# Containing risk

The ad-hoc proliferation of high-security biological labs must be controlled, and should be tied in more closely to broader research and public-health goals.

At the heart of any effort to control proliferation or improve safety, whether it concerns plutonium, toxic chemicals, disease agents or anything else, is the ability to take an inventory — to know with some confidence just how much material exists and where it is being kept. So it's disturbing to realize that authorities in many parts of the world often have only the vaguest notion of how many high-level biosecurity facilities exist on their territory, and that so few of them can document what stocks of deadly biological pathogens such facilities hold and exchange.

These labs have multiplied over the past decade as a consequence of concerns about bioterrorism following the terrorist attacks of 11 September 2001, and the appearance of public-health threats such as SARS and avian flu. New construction is already set to boost the number of biosafety-level-4 (BSL-4) labs in the European Union (EU) from 6 to 15 or more, and that in the United States from 7 to 13. BSL-4 is the highest level on the World Health Organization's rating scheme, and is required for labs that work with deadly, highly contagious pathogens, such as Ebola, Marburg, Nipah or Lassa viruses, for which no treatment exists.

Meanwhile, a September report by the US General Accounting Office (www.gao.gov/new.items/d09574.pdf) estimates that the number of American BSL-3 facilities has more than trebled, from around 415 in 2004 to 1,362 last year, But the report emphasizes the uncertainties. "The universe of BSL-3 and -4 laboratories is unknown," it states. "[T]here are likely other laboratories that we were unable to identify." Stocktaking in Europe is no better.

Building adequate high-containment infrastructure is important, in the sense that high-containment labs are essential for understanding exotic viruses, analysing samples from outbreaks and researching therapeutics and vaccines (see page 154). But the uncontrolled and uncoordinated nature of the growth of such facilities is problematic. Multiple agencies, countries and institutions too often pursue their own agendas, building biosecurity facilities with insufficient consideration for one another's plans or asking whether so much overall new lab capacity is needed — especially given the attendant risks of proliferation, accidental human infections and lab escapes.

The relevant authorities should undertake a much broader evaluation of such questions across institutions, focusing on how many and what sort of facilities are needed, if any at all; where they should be located; and how they could be coordinated to best suit the goals of broader research and public-health agendas. In the EU, for example, BSL-4 labs are currently concentrated in western and northern Europe. But they are lacking in southern and eastern Europe, where the BSL-4 class Crimean–Congo haemorrhagic fever virus is endemic.

> "The logic of having some sort of overarching oversight structure is sound."

Fortunately, such evaluation seems to be getting on the agenda. The EU has created a pilot for a proposed EU-wide network, the 'European research infrastructure on highly pathogenic agents' (ERINHA), which seeks to better coordinate the construction and operation of BSL-4 labs (see page 146). This is a step in the right direction, and should be complemented by a similar initiative for Europe's BSL-3 labs. Meanwhile, the General Accounting Office report recommends that the National Security Advisor, in consultation with federal agencies, designate a single body to evaluate all US high-containment labs. This recommendation should be taken forward.

In both cases, the logic of having some sort of overarching oversight structure is sound. Such structures could also provide much-needed mechanisms for the sharing of best safety and security practices and experience. Excessive oversight is undesirable. But the uncontrolled growth of labs, and the spate of instances of lax security in the United States in recent years (see Nature **461,** 577; 2009), indicate the need for greater independent oversight, and less laissez-faire. ∎

# No turning back

Spain should not use the recession as an excuse to stall plans to boost its scientific enterprise.

The past two decades have seen Spain transform itself from a scientific backwater into an internationally respected player in the research world. Much of that progress has occurred since the Socialist Party swept to power in 2004, pledging to turn Spain into an innovation economy (see Nature **451,** 1029; 2008).

During the Socialists' first term in office, for example, they doubled the science budget to just over €8 billion (US$12 billion), pushing it above 1.1% of the country's gross domestic product (GDP) and much closer to the European Union average of 1.8% of GDP. The party was re-elected in 2008, having pledged to cut bureaucracy and push funding for research to a target of 2% of GDP. Almost immediately it set up the Ministry of Science and Innovation, finally extracting science from under the purview of the education ministry. Cristina Garmendia, a molecular biologist who has founded several successful biotechnology companies, was appointed as head of the new ministry.

Since then, however, the momentum has been lost. Garmendia's political inexperience has shown. She was slow to build up a functioning ministry, and has not developed the necessary political clout to convince the government, now grappling with the global recession, to stick to its vision for science.

Granted, the government has bolstered financial support for the country's budding biotechnology and other high-tech industries. But

its draft budget for 2010, unveiled in September, proposed a cut of 45% for directly funded basic research. An outcry from the research community reduced that cut to 15%, and an extra 2.8% top-up for the science ministry is likely to emerge during parliamentary discussions. But this would still be a heavy blow to the country's research base.

Meanwhile, the government has yet to produce its much-heralded law for science. This was supposed to create an independent granting agency and reform the country's inflexible system of academic recruitment, under which university professors and government scientists are civil servants with an automatic right to employment until they retire. Dates for the law to be presented to parliament have been set and then withdrawn, apparently because some parts of the government do not want to exclude scientists from rules that apply to other government employees. Hiring new researchers continues to be a difficult and slow process, and it is almost impossible to offer a competitive package of salary and research money. The science ministry now says that the reform law will be presented to parliament before the end of the year, but the research community is losing faith that this will happen.

In the long-term, industry will be poorly served by a failure to develop and maintain a strong basic-research base. Spain is ill-advised to wed itself to the simplistic and outdated notion that a country can live on transferring knowledge while running down the knowledge generator. This is not a wise way to respond to the financial crisis.

Spain would do far better to emulate the commitments made last month by two other European nations as they too wrestle with the economic downturn. In Germany, a rich country with a near-stagnant economy, the centre-right government is cutting back public expenditure for 2010 everywhere except research and education, to which it is giving huge increases (see *Nature* **462,** 24; 2009). In Greece, a poor country with an economy in recession, the centre-left government says it will likewise cut public expenditure for 2010 everywhere except research and education, to which it is giving modest increases. The governments in both countries also plan to remove some of the red tape that restricts research.

Spain enjoyed one great period of intellectual brilliance in the early nineteenth century, referred to as its Silver Age. Until recently, Spanish scientists were optimistic that they were on their way to a second Silver Age. Now they joke that Spain is heading towards a Bronze Age. But they're not laughing. ∎

# A healthy get together

## The recently launched World Health Summit offers a rare chance for dialogue.

"Medicine is a social science," wrote the German pathologist Rudolf Virchow in 1848, "and politics is nothing more than medicine on a large scale."

That declaration by Virchow, a social activist and champion of scientific medicine who thought that medicine needed to "enter the political and social life" to fulfil its "great task", has lost little of its relevance today. Many of the problems of medicine are indeed social, and are often best addressed politically on a national or even global scale. Even as ageing populations are spending their last decades in need of medical care, expectations of medicine are rising with the promise of therapies based on genomic information or on stem cells, for example, that could be cripplingly expensive for health-care systems.

So it is, perhaps, fitting that the first World Health Summit, which was designed to address this divergence of the possible and the affordable in health care, took place on 15–18 October at the Charité in Berlin, the hospital where Virchow did his most important work.

The World Health Summit, which is planned as an annual event, grew out of the disappointment felt at the Charité when the G8 economic summit met in Germany in 2007 and — aside from a commitment to fight AIDS in Africa — failed to make health a priority on its agenda. Former Charité president Detlev Ganten teamed up with Axel Kahn, president of Paris Descartes University, and together they organized the M8 Alliance. It is a consortium of major medical institutions from eight leading countries that has taken responsibility for organizing this year's summit and all those that will follow.

The intent of the first health summit was not to produce specific decisions, but to provide a forum for dialogue and consensus-building among the world's many players in health — the full range of stakeholders who rarely have the opportunity to gather at the same time. Participants ranged from the researchers who develop high-tech therapies, to the politicians and funding agencies who support those developments, to the industries that will commercialize them, the physicians who will apply them and the health-care providers who will have to decide whether to pay for them.

In one session, for example, researchers working on regenerative medicine in the public and private sectors presented their latest thoughts on how to deliver stem cells into the body. And then they dashed eagerly into the next session where health-care providers presented their latest thoughts on reimbursing the cost of therapies based on regenerative medicine.

Such wide-ranging interchanges have been all too rare in the health arena, even as it becomes more important for medical researchers and health-care providers — and all stakeholders in between — to understand each others' potential and constraints. Health needs to be treated in the same comprehensive fashion that nations treat globalization, trade and finance. The first World Health Summit did not attract wide media attention. But over the years, it could and should develop some of the same political clout and visibility accorded to the G8.

The first summit also took place just weeks before the twentieth anniversary of the fall of the Berlin Wall, adding further historical resonance to the occasion. The Charité ended up in East Berlin after the Second World War. Although it had been a privileged research centre under the communist regime, the disappearance of the wall revealed the centre's infrastructure to be crumbling and its science outdated. It is testament to the success and confidence of reunited Germany that the Charité is now ranked as one of the top medical universities in the country, and is able to showcase such an important global initiative on its campus. That initiative is exactly the sort that Virchow would have championed, had he lived today. ∎

# RESEARCH HIGHLIGHTS



### Nerve repair

*Science* **326**, 818–823 (2009)

Researchers have slowed a fatal brain disease by inserting a gene into stem cells and then transplanting them into two patients. X-linked adrenoleukodystrophy (ALD) is a neurodegenerative disease caused by a mutation in the *ABCD1* gene. Patrick Aubourg of the French National Institute for Health and Medical Research (INSERM) in Paris and his colleagues engineered a virus to insert a functioning copy of the gene into the genome of the patients' bone-marrow stem cells. The team then depleted the patients of their bone marrow before infusing them with the repaired cells. Sixteen months later, neural degeneration had stopped in both patients. Two years after therapy, 18% of bone-marrow stem cells in one patient (four cells pictured) carried the working gene (cell with red dots).

For a longer story on this research, see go.nature.com/AZP3CC
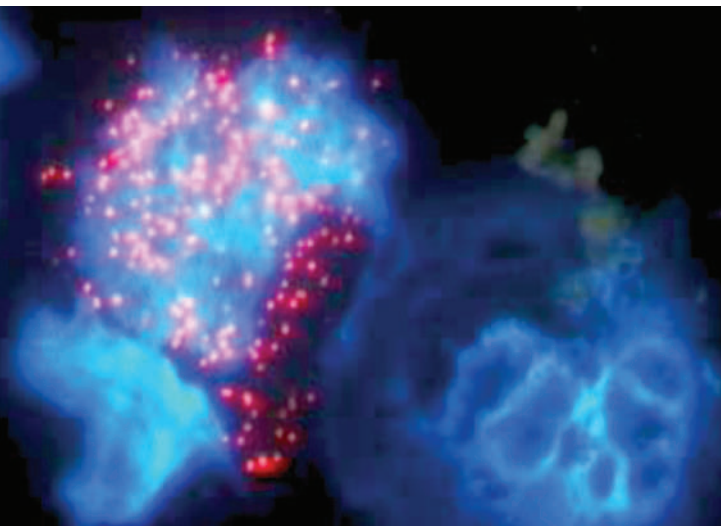
P. AUBOURG

---

## CLIMATE SCIENCE
### Volcano chills

*Geophys. Res. Lett.* 10.1029/2009GL040882 (2009)

Analyses of ice cores from Greenland and Antarctica have revealed a previously unknown tropical volcanic eruption that spewed huge amounts of ash and gas in 1809. The event resulted in a decade, starting in 1810, of the coldest temperatures recorded in the past 500 years.

Jihong Cole-Dai of South Dakota State University in Brookings and his colleagues identified unique isotopic signatures of volcanic sulphate in the core layers from 1810 and 1811 that point towards the eruption. Its exact location remains unknown. The researchers report that the release of sulphur gases, which form Sun-blocking aerosols, by this volcano and the famed 1815 Tambora eruption in Indonesia was sufficient to bring about the cold decade, which included the 'year without a summer' in 1816.

## LONGEVITY
### Sweet food, short life

*Cell Metab.* **10**, 379–391 (2009)

Nematode worms fed on a diet spiked with glucose die about 20% earlier than those consuming just the bacterium *Escherichia coli*.

Cynthia Kenyon and her colleagues at the University of California, San Francisco, found that dietary glucose inhibits the DAF-16 and HSF-1 proteins, which are known to lengthen nematode lifespan. This in turn lowers the activity of the gene *aqp-1*, which codes for a glycerol channel, suggesting that glucose shortens lifespan by affecting glycerol metabolism.

Worms consuming glycerol also died earlier. The authors think that the worms metabolize glucose into glycerol, which then initiates life-shortening processes.

## CLIMATE CHANGE
### Kilimanjaro's loss

*Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.0906029106 (2009)

Between 2000 and 2007, the area covered by the iconic glaciers of Mount Kilimanjaro shrank by 26%. In addition, the ice thinned rapidly, according to an analysis of aerial photographs and ground measurements by Lonnie Thompson of Ohio State University in Columbus and his colleagues.

One of the ice fields, the Furtwängler (pictured below), halved in thickness — by nearly five metres — between 2000 and 2009 at one drill site. The glaciers now cover just 15% of the area they covered in 1912. The



2000



2007

researchers predict that, probably owing to global warming, Africa's highest peak will be ice-free by 2033.

For a longer story on this research, see go.nature.com/IWMvvU

## NEUROSCIENCE
### Early stress marks genes

*Nature Neurosci.* doi:10.1038/nn.2436 (2009)

Changes in gene expression caused by factors other than variation in the DNA code — 'epigenetic' changes — are partly responsible for the mental and physical health problems often associated with stress in early life.

Dietmar Spengler and his colleagues at the Max Planck Institute of Psychiatry in Munich, Germany, stressed newborn mice by separating them from their mothers. As adults, the mice secreted abnormally high levels of the stress hormone corticosterone, were less able to cope with stressful situations and had memory impairments. They also had fewer methyl groups attached to the regulatory region for the gene that encodes the hormone vasopressin, a key player in the biochemical pathway that leads to corticosterone release. The reduced methylation resulted in a rise in vasopressin expression.

## ASTRONOMY
### Galaxies far, far away

*Astron. Astrophys.* doi:10.1051/0004-6361/200912299 (2009)

Astronomers have identified the most distant cluster of galaxies yet. The object, called JKCS 041, is 3.1 billion parsecs away, and existed when the Universe was just a quarter of its current age, say Stefano Andreon of the Brera Astronomical Observatory in Milan,

L. THOMPSON ET AL./PROC. NATL ACAD. SCI. USA

Italy, and his colleagues.

The team detected 16 of the brightest galaxies in the cluster, which consists of a gravitationally caged set of hundreds to thousands of galaxies. It is an extremely early example of the effect of gravity competing — and winning — against the dispersive effect of the Big Bang.

The researchers used both ground- and space-based telescopes, but required X-ray observations of hot gas between the individual galaxies to show that they are bound together in a cluster.

## BIOPHYSICS
# DNA stop and go

*Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.0907404106 (2009)
DNA polymerase enzymes that are responsible for DNA replication can work faster than previously thought.

Using a type of fluorescence spectroscopy, Jerrod Schwartz and Stephen Quake at Stanford University in California studied single polymerase molecules from the bacterium *Escherichia coli* in real time.

The enzyme pauses periodically as it travels along a strand of DNA synthesizing a partner strand, and the researchers measured its speed during periods of movement. They showed that the DNA polymerase Pol I(KF) has an intrinsic speed limit of 14–17 nucleotides per second, depending on the conditions — about ten times greater than estimates based on averages of all of its movements, including pauses. Another polymerase they looked at had a highly variable synthesis rate, ranging from 1 to about 50 nucleotides per second.

## ATMOSPHERIC SCIENCE
# Industrial UV shield

*Atmos. Chem. Phys.* **9,** 7737-7751 (2009)
Earth's natural sunscreen — the stratospheric ozone layer — has thinned during the past few decades because of the rise in atmospheric pollutants such as chlorofluorocarbons. This has allowed more ultraviolet (UV) radiation to reach many parts of the planet's surface since the 1970s. However, other forms of pollution have helped to shield Earth from UV rays.

Using an atmospheric radiation model, Gunnar Myhre of the University of Oslo and his colleagues found that, since 1750, pollutants such as sulphate, soot particles, sulphur dioxide and nitrogen dioxide have reduced the amount of UV light reaching some industrialized regions by as much as 20%. By scattering or absorbing UV light, such pollution may be masking some of the effects of ozone depletion, the authors say.

## NANOSCIENCE
# Release the goods

*J. Am. Chem. Soc.* doi:10.1021/ja9061085 (2009)
Silica nanoparticles can act as tiny containers to deliver drugs in response to enzymes, thanks to Chulhee Kim and his colleagues at Inha University in Incheon, South Korea.

The particles are pitted with pores that are capped with cyclodextrin molecules. The pores bear the cargo molecules — in the authors' experiments, fluorescent calcein dye.

The enzyme α-amylase, which increases in acute pancreatitis, was added to a buffer solution containing the particles. The enzyme degraded the cyclodextrin and calcein flooded out (schematic pictured below). The lipase enzyme had a similar effect, but when no enzyme was present, the pores remained closed and their contents locked away.

The system could have uses in drug delivery, diagnostics and imaging, the authors say.



AM. CHEM. SOC.
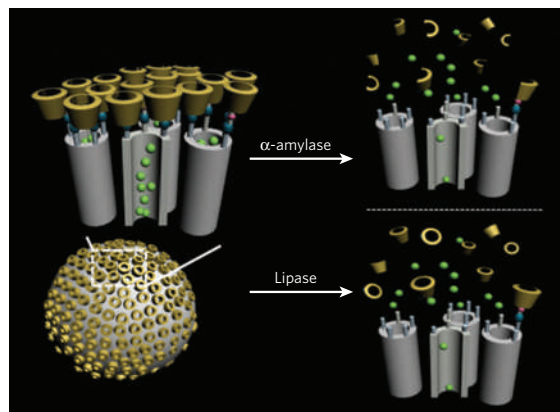
## GENOMICS
# Sequencing costs drop

*Science* doi:10.1126/science.1181498 (2009)
A team has sequenced three human genomes for US$4,400 each — at least ten times less than that achieved with other technologies.

Radoje Drmanac and Dennis Ballinger of Complete Genomics in Mountain View, California, and their collaborators developed a technique that chops DNA into fragments, makes many copies of these, rolls them up into 'nanoballs' and binds them to patterned silicon arrays. Nine-base-long probes of specific sequences — tagged with one of four different dyes, each corresponding to a 'letter' of the DNA code — are then added and bind to complementary DNA sequences with the help of synthetic 'adaptors'. Fluorescence imaging picks up the signal from the bound probes.

The technique keeps reagent use, and thus sequencing costs, low. It identified 94–98% of genetic variants when compared with established methods.

## JOURNAL CLUB

**Robert Blelloch**
**University of California, San Francisco**

**A computational biologist looks at how mRNA length changes during development.**

I am always amazed by how we start as a fertilized egg and develop into a complex, multicellular organism. This feat occurs despite the fact that the DNA in every cell — even the most specialized ones — remains, for the most part, unchanged.

One method of regulating gene activity in differentiated, or specialized, cells is through the messenger RNA (mRNA), the code of which is translated to make proteins. For example, proteins and other RNAs can bind to the untranslated regions (UTRs) at the 5' and 3' ends of mRNAs to regulate mRNA stability and translation.

The constitution of the 3' UTR itself can be regulated through alternative polyadenylation, whereby one of several possible UTR sites is cleaved, followed by the addition of adenosine-based molecules to its end. A broad shift in cleavage site choice — and thus 3' UTR length — during mammalian development was recently described by Bin Tian and his team at the University of Medicine and Dentistry of New Jersey in Newark (Z. Ji *et al. Proc. Natl Acad. Sci. USA* **106,** 7028–7033; 2009).

By analysing genomic data, they show that 3' UTRs generally get longer during development and cell differentiation. The authors further show that most of the genes in which 3' UTRs are lengthened are also those that are increasingly suppressed during differentiation, such as the genes for DNA replication and cell division.

These findings bring to the forefront an underappreciated mechanism of genetic regulation that is likely to be important for normal cell differentiation. It is fascinating how many steps of the central dogma (DNA to RNA to protein) are controlled. This seems to be how evolution has managed to take a relatively simple cell and multiply it to form the complex body plan of the human.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

# NEWS BRIEFING

## ● POLICY

**US–Muslim outreach:** Secretary of state Hillary Clinton has named three eminent US scientists as the nation's first science-and-technology envoys to Muslim-majority countries, where they will look to foster collaborative research programmes. The three are Bruce Alberts, editor-in-chief of *Science*; Elias Zerhouni, director of the US National Institutes of Health in 2002–08; and Ahmed Zewail, a Nobel laureate who serves on the President's Council of Advisors on Science and Technology.

**Montreal extension:** International delegates to the Montreal Protocol on Substances that Deplete the Ozone Layer concluded their meeting in Port Ghalib, Egypt, on 8 November without taking formal action to curb hydrofluorocarbons (HFCs), ozone-friendly refrigerants that could become a major contributor to global warming (see *Nature*, **461**, 1184; 2009). Some 41 countries joined North America, Europe, Micronesia and Mauritius in support of regulating HFCs as greenhouse gases. Delegates called on a technical committee to analyse alternatives to the chemicals in preparation for a potential decision next year.

**African science:** African governments could save millions of lives relatively cheaply by adopting health policies that are based on science, according to a report published on 9 November by seven of the continent's science academies. The report, which focuses on child and maternal health in sub-Saharan Africa, was unveiled at the fifth annual conference of the African Science Academy Development Initiative in Accra, Ghana. The first joint policy briefing to be produced by the initiative, it suggests that African governments are warming to the idea of taking advice from scientists. See go.nature.com/IysY5q for more.

## SPACE LIFT MARK ONE?

For the first time, a team has won some of the prize money at NASA's Space Elevator Games, held this year on 4–6 November at the Dryden Flight Research Center in southern California. LaserMotive, a company based in Seattle, Washington, used a laser to power a solar platform (pictured, prepared by team leader David Bashford) up a 900-metre suspended cable. It whizzed at speeds of 4 metres per second, winning US$900,000. To get the whole $2 million on offer required a climb at more than 5 metres per second. Aside from a powered climber, a space elevator needs a tether that is strong, thin and light enough to reach hundreds of kilometres into space. This component of the challenge, uncontested last week, remains unmet.



T. TSCHIDA, NASA/AP

**Nuclear shove:** The British government has declared ten sites as suitable for fast-track construction of nuclear-power stations. The 9 November announcement is part of a broader policy framework for meeting the nation's energy needs while keeping emissions of carbon dioxide in check. The plan also reaffirms a commitment to building a deep geological storage facility for nuclear waste. The United Kingdom produces some 15% of its electricity from nuclear power and hopes that will rise to 25% by 2025. See go.nature.com/4DCZdw for more.

**Advice row:** High-profile scientists in Britain have demanded that the government adopts new "principles for the treatment of independent scientific advice", following the sacking of government drugs adviser, David Nutt (see *Nature*, **462**, 11–12; 2009). Home secretary Alan Johnson, who sacked Nutt, has come under increasing pressure for his handling of the affair. As *Nature* went to press, the UK government said it would take steps to deal with scientists' concerns.

## ● BUSINESS

**Clean-tech takeover:** Electronics giant Panasonic attempted to catapult itself up the clean-technology league on 5 November by launching a ¥402-billion (US$4.5 billion) bid for a majority stake in Sanyo, another Japanese electronics firm. If successful, the takeover would combine two companies that together make most of the nickel-metal hydride batteries in hybrid cars, and a sizeable proportion of the world's lithium-ion rechargeable batteries; Sanyo is also a leading solar-cell manufacturer.

**R&D squeeze:** Pharmaceuticals firm Pfizer announced on 9 November that it would close 6 of its 20 research-and-development sites and cut research at some remaining centres to leave five main hubs, following its US$67-billion acquisition of fellow drug maker

## SOUND BITES

**"If you look at our international security system, it is in tatters."**

Mohamed ElBaradei, who leaves the International Atomic Energy Agency, the United Nations' nuclear watchdog, at the end of this month. After 12 years in charge, he is not impressed by efforts to curb nuclear proliferation.

Wyeth (see *Nature* **457,** 520–521; 2009). The company did not say how many jobs would be lost directly from these shutdowns; it has previously said it would cut 15% (around 19,500 jobs) of the combined firms' workforce.

**Monsanto in China:** The biotech seed giant Monsanto formally opened its first biotech research centre in China on 4 November, adding to similar facilities in the United States, Brazil and India. The centre, in Beijing, will concentrate on early-stage bioinformatics and genomics research and serve as a base for collaborations with Chinese scientists. The facility will focus on Monsanto's core crops of maize (corn), soya beans and cotton.

**HIV collaboration:** Drug companies Pfizer and GlaxoSmithKline formally launched their joint HIV venture, ViiV Healthcare, which will develop combination drugs using the research and market portfolios of its parent companies. The move was first announced in April (see *Nature* **458,** 950–951; 2009).

### ● RESEARCH

**Industry ties:** The proportion of US life scientists whose research is directly funded by industry is dropping, according to a 2007 survey published last week (D. E. Zinner *et al. Health Affairs* **28,** 1814–1825; 2009). Of more than 2,000 researchers surveyed, just over half said they had some kind of relationship with industry — mostly serving as consultants, paid speakers or members of scientific advisory boards. But only 20% of research faculty received industry funding, compared with 28% in 1995, the last time a similar poll was conducted.

**Theoretical physicist dies:** Vitaly Ginzburg, who won the 2003 Nobel Prize in Physics, died on 8 November aged 93. Working at the P. N. Lebedev Physical Institute in Moscow, Ginzburg (pictured below) helped to develop the Soviet hydrogen bomb in the 1940s, a project whose importance shielded him from the worst of Stalin's purges. "I was saved by the hydrogen bomb," he said, after sharing the Nobel award for his work on the theory of superconductivity.

**University competition:** The UK government plans to concentrate public research funds on the nation's top few universities, according to a policy document published on 3 November. The report says that funds will be diverted towards institutions providing science and engineering degrees; researchers from different institutions might also have to collaborate, rather than compete with each other, for scarce funds, it adds.

### ● FUNDING

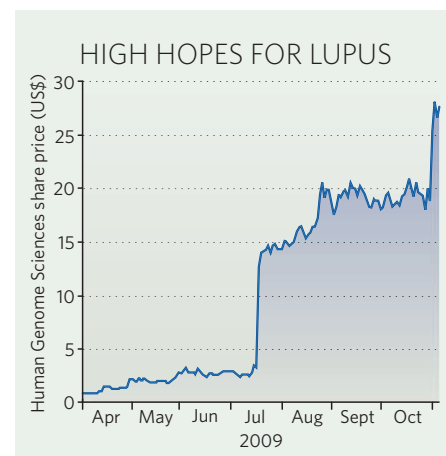**Personal grants:** The UK Wellcome Trust, the country's largest non-governmental funder of biomedical research, will from autumn 2010 offer long-term research grants focused on individual scientists rather than on specific project applications. The move was in part a response to researchers' complaints about grant sizes and short funding periods. See page 144 for more.

**"The bird escaped unharmed but lost its bread."**

A statement from CERN, the European particle-physics lab near Geneva, Switzerland, on the cause of a short circuit that disrupted the Large Hadron Collider last week. An errant bird dropped a baguette on a piece of electrical equipment. See go.nature.com/mDzPFV for more.

## THE WEEK AHEAD

**11 NOVEMBER/SPIRIT SOL 2083**
The NASA rover Spirit, stuck in soft soil on Mars for six months, is due to begin its escape. The recovery effort could start as early as 11 November, on the rover's 2083rd Martian day.
▶ go.nature.com/Jk7EF5

**12–13 NOVEMBER**
The Royal Society, in London, hosts a meeting on genetics and the causes of evolution, summarizing "150 years of progress since Darwin".
▶ go.nature.com/gUhoxi

**16–18 NOVEMBER**
The World Summit on Food Security meets in Rome, Italy, under the auspices of the United Nations' Food and Agriculture Organization (FAO). Ahead of the meeting, the FAO warned that food prices in some poor countries remained stubbornly high, above pre-economic-crisis levels.
▶ go.nature.com/yqpqaf

**18–19 NOVEMBER**
In Cadarache, France, the council of ITER meet to discuss new cost estimates for the multibillion-euro project, which aims to prove controlled nuclear fusion.
▶ go.nature.com/Hv6smk

# BUSINESS WATCH

Besides steroids and other general treatments for autoimmune diseases, it has been half a century since a drug was approved for systemic lupus erythematosus. Lupus affects mainly younger women, causing ailments from arthritis to kidney failure, and sometimes results in death.

But last week, Human Genome Sciences (HGS), a pharmaceutical company in Rockville, Maryland, announced positive results from a second, late-stage, clinical trial of its monoclonal antibody Benlysta (belimumab), developed in partnership with drug giant GlaxoSmithKline. Benlysta specifically targets a protein that goes into overdrive in lupus, causing the body to produce antibodies that attack its own tissues.

HGS's share price rose 35% on 2 November — a single-day gain bettered only in July when its first phase III clinical trial results were announced (see chart). The news "opens the door to one of the more exciting 'new drug' stories in biotechnology in recent years. The drug defied the axiom that 'nothing ever works in lupus'", wrote Geoff Porges, an analyst at New York-based financial research firm Sanford C. Bernstein, in an investor note. Porges thinks that both the US Food and Drug Administration and the European Medicines Agency will approve the drug, and — as there are no well-placed competitors — global revenue from Benlysta could exceed US$1.5 billion in 2012.

### HIGH HOPES FOR LUPUS

Human Genome Sciences share price (US$) — 2009: Apr, May, Jun, Jul, Aug, Sept, Oct

# NEWS

# End of the road for Copenhagen?

Expectations are dropping as December's UN climate talks get closer.
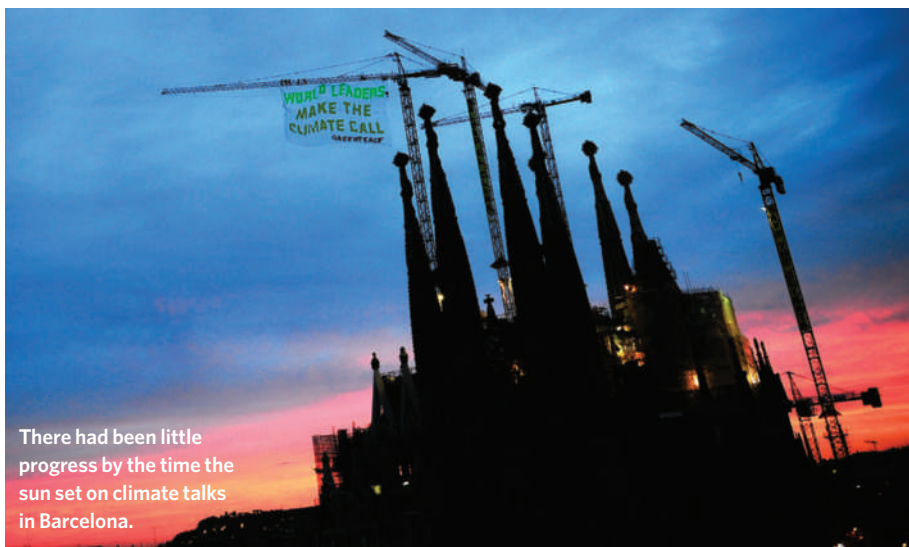
**BARCELONA**

In the final round of talks before the Copenhagen climate summit in December, negotiators have tried to map out a path ahead in the likely event that a fully binding international treaty cannot be signed this year.

Delegates at the 2–6 November meeting in Barcelona made progress in winnowing out the text of a treaty that they hope will succeed the Kyoto Protocol, which expires in 2012. But opposing positions on the core issue of reducing greenhouse-gas emissions seem to have become more entrenched.

On the last day of the conference, Ambassador Lumumba Di-Aping of Sudan again explained the dangers of planning for a warmer climate rather than acting to prevent one.

By refusing to take aggressive measures to curb greenhouse-gas emissions, he said, industrialized nations are condemning the poorest and most vulnerable people to disasters that no amount of money or technology can prevent. Lumumba, who as chair of the G77 represents most of the developing world and a majority of countries in the talks, had delivered the same message each day in Barcelona. This time his deep monotone voice echoed throughout a room full of empty chairs; the handful of journalists present asked no questions.

Despite repeated criticism from Lumumba and others, the European Union and the United States made it clear that there is little room for negotiation on their commitments



There had been little progress by the time the sun set on climate talks in Barcelona.

D. RAMOS/AP PHOTO

to cut emissions by 2020, which fall far short of the 40–45% below 1990 levels that developing countries are demanding.

Indeed, a key question in Copenhagen will be whether poor countries are willing to accept less ambitious cuts in exchange for additional resources to help them cope with a warmer world and institute sustainable development plans (see 'Key issues for Copenhagen').

Lumumba says the answer is no. If industrialized countries refuse to step up their commitments, there will be no deal in Copenhagen, he says, and "we simply accept that we have failed".

Such anger and distrust is widespread among representatives of developing countries, many

of whom feel bullied in the negotiations and distrust carbon markets as another money-maker for Western industrialists. The situation boiled over on the first day of the talks in Barcelona, when one of two key negotiating tracks was suspended for a day after African countries refused to talk about anything except for emissions targets for developing countries.

But the G77 is a diverse group that fractures on many questions, and Lumumba doesn't speak for all on this issue. South Africa's chief negotiator Alfred Wills talks tough on emissions, but acknowledges that additional money and resources could soften opposition from developing countries. "That's part of the negotiations," he says.

**Willing to negotiate**

European Commission chief negotiator Artur Runge-Metzger said that the EU conditional offer to cut emissions by 30% below 1990 levels by 2020 is within the 25–40% range cited by the Intergovernmental Panel on Climate Change (IPCC) as likely to prevent more than 2 °C of warming (see graph). Pressed by reporters on repeated statements by the G77 that a 40% cut is non-negotiable, Runge-Metzger didn't budge. If the G77 weren't willing to negotiate, he told reporters, "they wouldn't be here".

US negotiator Jonathan Pershing said the commitments being offered by developed countries today are the result of intense national debates, which makes them more robust but also means they are unlikely to
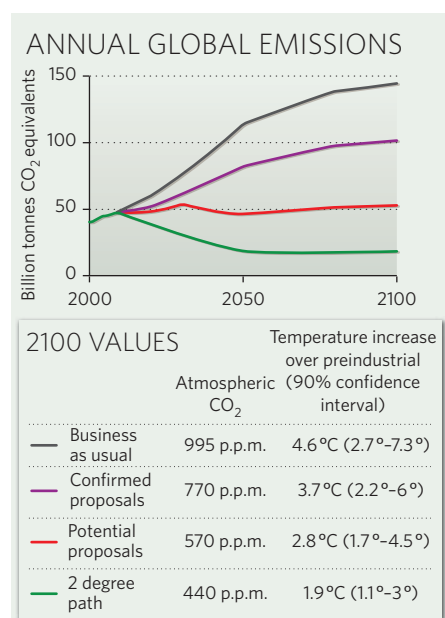
## Key issues for Copenhagen

**Preparing 'Plan B'.** Although nobody actually calls it that, leading negotiators and United Nations officials have outlined proposals on what to do if the talks in Copenhagen fail to deliver a full treaty — aim for political agreement on the major issues and hope to fill in the details in early 2010.

**One treaty or two?** Developing countries want to maintain the Kyoto Protocol and create an additional agreement that includes the

United States, rather than creating a new protocol that covers everybody.

**Regional centres of excellence.** Many negotiators are optimistic about a proposal to establish regional centres to focus on technology transfer and capacity building in developing countries to help them cope with climate change.

**Deforestation.** Direct public financing could help tropical countries build up

their science and regulatory capacity to monitor deforestation, before being allowed to participate in carbon trading.

**Adaptation financing.** Agreement is building on creating a start-up fund, probably administered under the United Nations climate framework, to help developing nations adapt to climate change. Disputes remain over how much cash is needed.                **J.T.**

change over the coming month. He went on to criticize many developing countries for trying to avoid commitments to reduce their own potential growth in emissions.

Pershing also sought to head off rampant speculation that the United States would not be ready to sign a treaty in Copenhagen, given that several alternative climate bills are still pending in Congress. Those bills offer 17–20% emissions reduction below 2005 levels by 2020, he said, enough for the world to be able to judge the US position.

However, the economic growth of the 1990s means that such cuts would equate to a reduction of just a few per cent below 1990 levels. That's not enough, says the Alliance of Small Island States, which is calling for industrialized nations to promise a 45% emissions reduction by 2020 compared with 1990 levels. The alliance



CLIMATE INTERACTIVE

### ANNUAL GLOBAL EMISSIONS

| 2100 VALUES | Atmospheric $CO_2$ | Temperature increase over preindustrial (90% confidence interval) |
| --- | --- | --- |
| Business as usual | 995 p.p.m. | 4.6 °C (2.7°–7.3°) |
| Confirmed proposals | 770 p.p.m. | 3.7 °C (2.2°–6°) |
| Potential proposals | 570 p.p.m. | 2.8 °C (1.7°–4.5°) |
| 2 degree path | 440 p.p.m. | 1.9 °C (1.1°–3°) |

points out that since the last IPCC report in 2007, new research has suggested that there will be even greater impacts on their nations than previously estimated, including flooding, erosion and an upsurge in extreme weather.

Alden Meyer, director of strategy and policy for the Union of Concerned Scientists, in Washington DC, says developed-country commitments are indeed less ambitious than they should be, but nonetheless represent a solid base that can be built on in the future. Like many others, he thinks that countries are likely to converge on some kind of agreement in Copenhagen, even if negotiators need to continue working out the details next year. "We probably aren't going to get what we want in Copenhagen," he says. "But can we get what we need? I think we can." ∎

**Jeff Tollefson**

# Wellcome Trust makes it personal in funding revamp

The Wellcome Trust, the UK's largest non-governmental funder of biomedical research, is giving its grant scheme a major overhaul.

The trust plans to stop asking researchers to submit extensive applications detailing their proposed work over the three- or five-year period of a grant. Instead, it will award longer-term funding focused on individual researchers rather than on specific research projects. "The best way to fund science is to fund people," says Mark Walport, director of the Wellcome Trust.

The new Wellcome Trust Investigator Awards for junior faculty members, and the Senior Investigator Awards for more experienced scientists, will assess researchers and their ideas primarily through an interview process, with funding awarded for five to seven years. Researchers will still have to submit written applications, but "we don't want people to focus on the precise details of how the particular experiment will be done", says Walport.

Walport hopes that the shift in emphasis will allow researchers to focus on presenting a big scientific question and why it is important, with applicants grilled on the details of how they plan to tackle the problem during interview.

The investigator awards build on the trust's existing fellowships, which, like most funders' fellowships, are not given to researchers as a supplement to existing salaried academic positions. "When we interview people for fellowships, they have an opportunity to present their case, and one can have a good exploration as to whether they have a spark and original ideas. We want to extend that idea to people with salaried positions to give people a decent quantum of funding," says Walport. Winners will have the freedom to spend their funding on any research question they want, and will not be limited to the ideas they proposed in their application, he says.

The trust has yet to agree the exact amount of money it will commit to the investigator awards but expects it to be about the same amount that it currently

awards to project and programme grants, which totalled around £110 million (US$180 million) in 2008. The first round of applications for the investigator awards is expected in autumn 2010, with funding to commence at the beginning of 2011. The move is in part a response to complaints from scientists about the size of grants and the short funding periods generally set by funding bodies in the United Kingdom, says Walport.

Given the more extensive nature of the investigator awards, fewer scientists will be supported than under the current scheme. But Jim Smith, director of the National Institute for Medical Research in Mill Hill, UK, and a previous winner of several Wellcome Trust grants, says that "it is better to fund one person properly than two or three people inadequately". He adds that if the trust's scheme proves to be successful, other funders should consider similar approaches.

The trust says that it is not aware of any other funding body in the United Kingdom that offers the same style of funding. But in the United States, the Howard Hughes Medical Institute (HHMI), in Chevy Chase, Maryland, has operated a similar scheme since the late 1980s. It now runs funding rounds roughly every three years, awarding about $60 million in total based on short written applications outlining a broad scientific question, and then on a more detailed interview. Winners receive awards lasting five years, and are given total freedom to spend the money on any medical research question they wish to pursue. "Traditional grants don't look so favourably on researchers if they venture outside the area they said they were going to tackle," says Jack Dixon, HHMI vice-president and chief scientific officer, adding that giving researchers the freedom to pursue discoveries made along the way produces "better science". The approach is more honest, agrees Smith, because it acknowledges that science has often moved on between the grant application being made and work actually beginning. ∎

**Natasha Gilbert**

> "The best way to fund science is to fund people."

# European biosafety labs set to grow

**Bioterrorism and emerging diseases spur building boom, although some question the need for more facilities.**

A rash of new maximum-security biosafety-level-4 (BSL-4) labs are being built in the European Union (EU), and even more are planned, but some scientists are arguing that the bloc already has more facilities than it needs.

The issue is currently being tackled by a three-year study, costing €5 million (US$7.5 million), which looks set to determine the future of the European BSL-4 landscape. The study, funded by the European Commission, is the initial phase of a project called 'European research infrastructure on highly pathogenic agents' (ERINHA), which will bring all EU BSL-4 labs into a single network.

The project involves the main laboratories and national governments, and aims to work out a coherent scientific and policy case for future BSL-4 needs, thereby helping to avoid duplication, says Hervé Raoul, director of the Jean Meriéux-Inserm BSL-4 lab in Lyon, France, and coordinator of ERINHA. It would also introduce pan-European guidelines on best safety and security practices, and enable researchers from any member state to use facilities anywhere in the EU.

ERINHA is part of a broader EU initiative, the European Strategy Forum on Research Infrastructures (ESFRI), that was set up to improve the coordination of Europe's large research infrastructures (see *Nature* **450,** 586; 2007). It foresees costs of some €174 million for construction and upgrading of BSL-4 labs, and €24 million annually for operating expenses, with most funding coming from member states.

The EU currently has six operational BSL-4 labs in four countries — Britain, Germany, Sweden and France — with at least eight new facilities being built, or under discussion, in Italy, Germany, the Netherlands, Switzerland and Austria (see map), although the Italian facility will replace the existing one there. That expansion mirrors a similar boom in the United States, where new construction will increase the number of BSL-4 labs from 7 to 13 (see *Nature* **461,** 577; 2009).

Scientists contacted by *Nature* agree that there are benefits from greater collaboration between centres and from refurbishing older labs. Bringing the BSL-4 labs together in a network should also help to improve the response to health threats at a pan-European level, says Carla Nisii, a researcher at the National Institute of Infectious Diseases in Rome. The project also
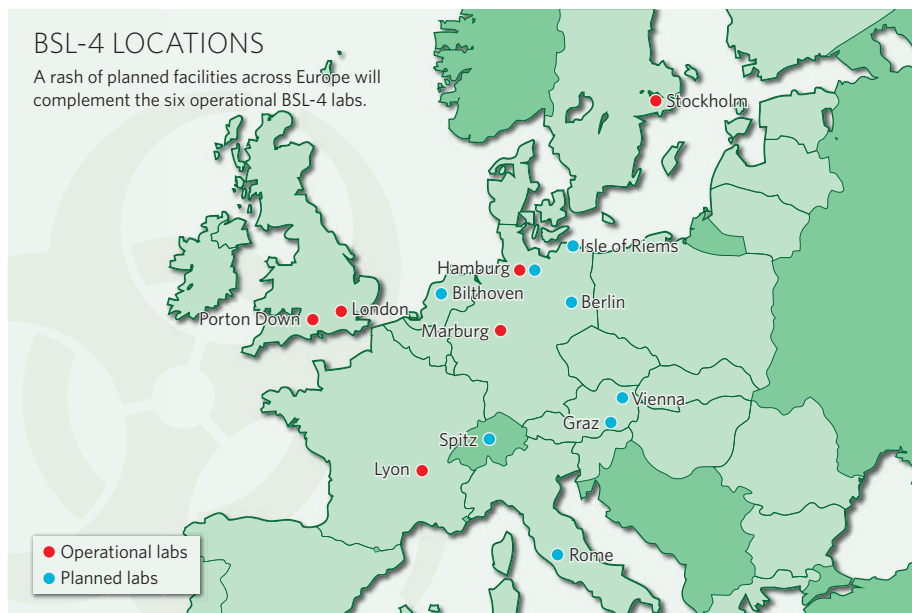


**Containing risk: a project to link up all of Europe's highest-security labs is being planned.**

aims to build laboratories in Eastern Europe.

But some are sceptical about the need for more European facilities. "The existing facilities are already more than sufficient," says one high-security-lab researcher who requested anonymity. "I'm not so sure more infrastructure is necessary," agrees Stephan Günther, head of the Bernhard Nocht Institute for Tropical Medicine in Hamburg, Germany, which has a decades-old BSL-4 lab and is preparing to open the doors of a new one. Günther's lab specializes in research on exotic and dangerous viruses, as do most of the older BSL-4 labs in Europe.

Günther suggests that the current boom in BSL-4 labs has been largely prompted by heightened concerns over bioterrorism following the terrorist attacks of 11 September 2001. "Before 11 September there was no interest in

## BSL-4 LOCATIONS

A rash of planned facilities across Europe will complement the six operational BSL-4 labs.



- Stockholm
- Isle of Riems
- Hamburg
- Bilthoven
- Berlin
- Porton Down
- London
- Marburg
- Vienna
- Graz
- Spitz
- Lyon
- Rome

● Operational labs
● Planned labs

## Risk ratings

**BSL-1:** For basic teaching or research use, handling organisms unlikely to cause disease.

**BSL-2:** For health and diagnostics services, handling pathogens that cause treatable diseases, such as rabies virus or the 2009 pandemic H1N1 flu virus. Researchers wear protective clothing and use biosafety cabinets.

**BSL-3:** For specialized diagnostics or research on pathogens that cause serious but not highly contagiousdiseases, such as SARS coronavirus. Sealed labs have negative air pressure and safety cabinets for all work.

**BSL-4:** Dangerous-pathogen units for research on exotic agents that causes serious and highly contagious diseases, such as Ebola virus. Lab entry via airlocks, exit via showers. Researchers work in glove-box safety cabinets with air-supplied positive-pressure 'spacesuits', and are in constant radio contact.

**KEEPING SYNTHETIC
DNA AWAY FROM
TERRORISTS**
Gene-makers suggest
security standards.
go.nature.com/AytMyG
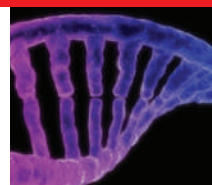
INGRAM PUBLISHING

# Call to boost isotope supplies

The US Department of Energy should build two dedicated isotope-production facilities, costing about $65 million in total, to solve worsening supply problems for researchers in medicine, physical sciences and national security. That's the conclusion from a panel convened by the energy department's Nuclear Science Advisory Committee (NSAC), which approved the panel's report on the state of the US isotope programme on 5 November.

The programme supplies researchers with isotopes that are not readily available from commercial suppliers, and is tiny compared with the vast market for routinely used medical isotopes, such as technetium-99m — which itself is still beset with ongoing supply problems (see *Nature* 460, 312–313; 2009).

Despite the programme's small size — its 2008 budget was just $32 million — its products are essential to a wide array of research fields. But fragmented and ageing production facilities at the energy department have struggled to keep up with the variety and pace of demands. So last year the department commissioned the NSAC to identify the most important research isotopes and to come up with ways to alleviate supply fluctuations.

*"Other priorities can sometimes bump research-isotope production to the back of the line."*

The committee concluded that a group of isotopes with potential for use in medical therapy were the most critical. These isotopes, including actinium-225, emit α-particles that have high energies but low velocities, which means that they are effective at killing tumours without damaging healthy tissue. But shortages are holding up clinical trials, says Roy Brown, who was an industry representative on the report and is director of federal affairs for the Council on Radionuclides and Radiopharmaceuticals, which represents US and Canadian isotope manufacturers. Other important medical isotopes include arsenic-76, used in palliative care for bone pain, and palladium-103, implanted as seeds into prostate glands to kill cancers.

But isotopes are used for more than just medicine. Physicists want californium-252 so they can split its heavy nucleus to make beams of smaller, rare isotopes, useful for frontier experiments in nuclear physics. NASA wants better supplies of plutonium-238 as a thermal heat source for long-lived planetary probes. And germanium-76 is needed for decay experiments that test whether neutrinos are their own anti-particles, which could help explain why the Universe is dominated by matter rather than anti-matter.

Isotopes used in national security often take precedence over other research needs — especially in the case of helium-3, which is being used in neutron detectors at ports to spot smuggled plutonium. But this has pushed up prices for researchers who want helium-3, used in many of the ultra-low-temperature refrigeration systems needed, for example, to study the super-cooled clouds of atoms known as Bose–Einstein condensates.

The report says that supplies of many of these isotopes could be much improved by building two new facilities. One would be a electromagnetic separator to enrich certain rare isotopes; the other would be an accelerator, which could collide different particles to create isotopes that are not found naturally. Donald Geesaman, a physicist at Argonne National Laboratory in Illinois who co-chaired the report committee, says that the separator could be built for $25 million and the accelerator might cost $40 million.

Research isotopes managed by the programme currently come from three places: accelerators at Brookhaven National Laboratory in Upton, New York, and Los Alamos National Laboratory in New Mexico, as well as from a nuclear reactor at Oak Ridge National Laboratory in Tennessee. But producing research isotopes is a secondary task for these facilities, and other priorities can sometimes bump isotope production to the back of the line, says Brown.

Jehanne Gillo, who directs the isotope programme for the energy department, says that the report comes too late to be included in the department's budget request for fiscal year 2011, but could be used to compete for money against other requirements in 2012.
Eric Hand

---

all these diseases; it was only tropical-health researchers who were interested in them," he says.

"Bioterrorism is taken into account, of course," says Raoul. But the primary motivation is to help the EU prepare for emerging diseases such as SARS. BSL-4 labs are a vital part of the standard response to an outbreak, responsible for isolating and characterizing the pathogen, developing diagnostic techniques and sometimes working on potential therapies. "We are seeing roughly one new emerging or re-emerging pathogen per year, while pathogens are also changing their geographical ranges, and travel is resulting in more imported cases of exotic diseases," says Philip Luton, a scientist who is now head of business development and spokesman at the UK Health Protection Agency's Centre for Emergency Preparedness and Response at the Porton Down military establishment near Salisbury.

Although Günther is sceptical about the need for more labs, he thinks that the ERINHA project would certainly help existing facilities to upgrade. Whereas national governments may fund the construction costs of BSL-4 labs, they are often unwilling to cover the running costs of such labs in the long term, he says, a problem that a pan-European network could help to address. The running costs of a BSL-4 lab are much higher than those of a typical virology lab, says Raoul. Maintenance of his own Lyon lab runs to €1.5 million annually, on top of €1 million in salaries for the core support staff who assist visiting researchers.

The best way to ensure sustained operational funding is to make a well-argued case at the European level and to get buy-in from national decision-makers, argues Raoul. "We are making it clear that to construct a BSL-4 lab without taking into account its subsequent running costs is suicidal." ■

**Declan Butler**
**For more, see Editorial, page 137, and News Feature, page 154.**

**HAVE YOUR SAY**
Comment on any of our
News stories, online.
**www.nature.com/news**

# China moves to help high-tech firms

Two initiatives to boost investment in high-tech companies in China have received a mixed response from entrepreneurs.

Trading in shares in the first 28 companies on ChiNext, a new stock exchange for innovation-oriented companies, started on 30 October. On the same day, the government's National Development and Reform Commission (NDRC) announced that it was creating 20 venture-capital funds that would be worth 9 billion renminbi (US$1.3 billion), a significant proportion of the country's venture-capital market (see graphic).

ChiNext, part of the Shenzhen Stock Exchange (SZSE) in Shenzhen, is being called China's answer to the NASDAQ, the US market considered the natural home for high-tech businesses. On ChiNext's opening day, share prices on the SZSE grew so quickly that regulators were forced to suspend trading, fearing dangerous inflation of stock.

"[ChiNext] will speed up the establishment of spin-offs and provide an effective financing channel to commercialize technologies," says Li Yang, a project manager at the International Technology Transfer Center at Tsinghua University in Beijing, which has one of the strongest university intellectual-property portfolios in China.

Lawrence Lau, an economist and vice-chancellor of the Chinese University of Hong Kong, agrees that ChiNext is a positive development because it provides a convenient way to make money by issuing stocks in small Chinese companies, a channel that has not previously existed (see *Nature* **437,** 1087; 2005).

Greg Scott, founder of the Shanghai-based



**Shares in companies quoted on the new ChiNext stock exchange soared on the opening day.**

consulting firm ChinaBio, points out that four of the five pharmaceutical companies listed on ChiNext are producing or planning to produce novel drugs. Until now, he says, most successful Chinese pharmaceutical companies have focused on medical devices or generic versions of existing drugs. "Seeing so many [innovative companies] in the drug space surprises me," adds Scott.

Tom Baruch, founder of CMEA Capital, based in San Francisco, California, says that the new market makes foreign investment in

Chinese companies more likely. "I am sure there will be a great deal of foreign interest in the stocks listed on ChiNext," agrees Lau, "but China does not really need more foreign capital."

Some critics point out, however, that companies must clear significant hurdles before they can be listed. Businesses must show a cumulative profit of 10 million renminbi over the past two years, or a profit of 5 million renminbi in the past year with revenue of 50 million renminbi and a two-year revenue growth rate of at least 30%.

That's difficult for most small, innovative companies in need of funding, says Qiang Yu, chief executive of Shanghai Ambrosia Pharmaceuticals. "The first companies [quoted on ChiNext] are not high tech and high risk; they are quite mature and profitable."

Yu and Mingdong Zhou, founder and chief executive of Zensun (Shanghai) Science & Technology Ltd, have both struggled to raise funds to develop drug candidates, and they expect little help from the new market. "Domestic innovations are not the target of ChiNext," says Zhou. "There is virtually no financial mechanism to support innovative drug development in China, except for basic research." Clinical trials are particularly difficult to finance, he adds.
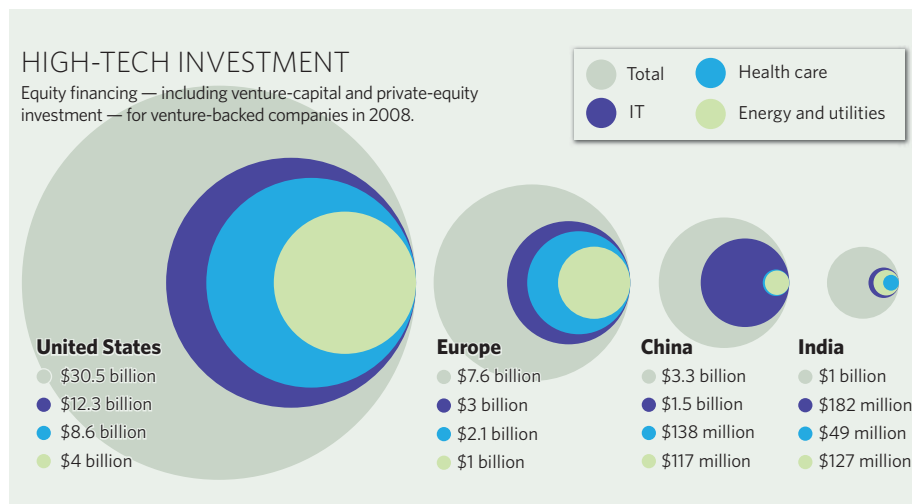
### No sea change

Venture capital for small pharmaceutical and biotech companies has long been hard to come by in China, and some question whether the government's new fund will change that.

In the NDRC scheme, central government is providing 1 billion renminbi, and seven local governments are chipping in with a total of 1.2 billion renminbi. The rest of the 9-billion-renminbi fund is expected to come from private sources.

But little of the funding is likely to reach biotechs, predicts Scott, because the scheme relies so heavily on private investors, who are usually wary of long-term drug development.

"These funds are for a variety of technologies, including IT, electronics, energy and environment. I think life sciences will see very little — private investors much prefer the faster returns provided by the tech industries," he explains. "Only about 5% of venture-capital funding in China goes into life sciences. If you extrapolate that to the new funds, a total of about US$64 million would end up in life sciences — helpful, but not a sea change."
**David Cyranoski**

## HIGH-TECH INVESTMENT
Equity financing — including venture-capital and private-equity investment — for venture-backed companies in 2008.



Total
Health care
IT
Energy and utilities

**United States**
- $30.5 billion
- $12.3 billion
- $8.6 billion
- $4 billion

**Europe**
- $7.6 billion
- $3 billion
- $2.1 billion
- $1 billion

**China**
- $3.3 billion
- $1.5 billion
- $138 million
- $117 million

**India**
- $1 billion
- $182 million
- $49 million
- $127 million

**US HABITAT RULE THREATENS SPECIES**
Call for change to definition of 'endangered'.
**go.nature.com/XSLFoC**

## SNAPSHOT
# Glider eavesdrops on whales

Quietly slipping to a depth of 1,000 metres, an undersea glider is recording whale song off the coast of Hawaii in unprecedented detail.

It is the first acoustic-equipped glider to be deployed to this depth in the ocean to target a specific marine mammal. Whales click or vocalize to communicate and to find food, and use echolocation to navigate, but surface acoustic devices typically can't record their sounds.

Since 27 October, the glider has made more than 60 dives, each lasting about 6 hours, and is due to be retrieved on 17 November. It will collect data on beaked whales, which seem particularly sensitive to man-made noise; several strandings of these whales have been associated with military sonar usage (see *Nature* **425,** 575–576; 2003).

"We believe we have identified beaked whales," says Dave Mellinger of Oregon State University, Corvallis, part of the project team. "It was pretty exciting. You work a couple of years on a project, hope it will succeed, but you don't know until the equipment is wet."

The US Office of Naval Research is funding the $1.5-million project, which builds on more than a decade of using autonomous gliders to study ocean temperatures and currents. This glider is steered by an internal computer on a preprogrammed course, travelling at about 0.25 metres per second, and is expected to collect half a terabyte of data over the course of its cruise.

**Rex Dalton**

# Britain sets up defence advisory group

The British government has recruited a group of academics to tackle tricky scientific problems related to defence, *Nature* has learned.

The programme is similar to a group known as the JASONs, which the US government has consulted on technical issues since the 1960s. "You hear a lot about the JASONs and how much credibility they have in the United States," says Mark Welland, the UK Ministry of Defence's chief scientific adviser. Britain needs a similarly "fast-moving, free-floating entity", he says.

Scientific advice is frequently sought in Britain, but on security-related issues the advice usually comes from inside the government. Scientists at government labs such as the Atomic Weapons Establishment in Aldermaston are consulted on sensitive topics, in part because academic researchers lack the necessary security clearances.

The situation contrasts with that in the United States, where academic scientists routinely travel in and out of classified government laboratories, often maintaining their clearances after they have left a lab. The JASONs, a semi-secretive

group of roughly 30 academics, typically meet over the summer to look at technical problems faced by the Pentagon. During the cold war, the group was considered indispensable for its work on problems such as submarine detection.

It was a model that Britain lacked but


Can 'Newton's Apples' help tackle roadside bombs?

needed, Welland says. So in April, Welland and John Beddington, the government's chief scientific adviser, assembled 11 academics in an attempt to duplicate the success of the JASONs. The group was tasked with looking at ways to improve radiation detection at

the nation's ports to prevent terrorism and smuggling of nuclear material.

John Hassard, a physicist at Imperial College London, is one the researchers chosen to participate in the project's wide-ranging discussions. "At some points we were talking about some pretty far-out ideas, such as extrasensory perception and gravity waves," he says. But the group settled on less-radical solutions involving plasma physics that Beddington says are now being considered for funding.

Beddington and Welland say that they are now planning to hand-pick a second group to look at improvised explosive devices. Such home-made bombs are commonly used by insurgents in Iraq and Afghanistan, and Welland says that the government wants to get fresh ideas about how to deal with the devices at every stage, from their production to their detection and deactivation. "It's a high priority for the Ministry of Defence so it is an obvious area for a group such as this to really get to grips with," Welland says. "What we don't have yet is a name for this group," he adds. "I suggested that they be called Newton's Apples."

**Geoff Brumfiel**

# Report row ousts top Indian scientist

Ruckus over call for reform at national science agency raises fears for attracting expatriate talent.

**BANGALORE**

The first appointment in a scheme to recruit expatriate scientists to senior positions in the Council of Scientific and Industrial Research (CSIR) — India's largest science agency — seems to have misfired badly.

A US scientist of Indian origin has been dismissed just five months after he was offered the position of 'outstanding scientist' and tasked with helping to commercialize technologies developed at CSIR institutes.

Shiva Ayyadurai, an entrepreneur inventor and Fulbright Scholar with four degrees from the Massachusetts Institute of Technology in Cambridge, was the first scientist to be appointed under the CSIR scheme to recruit about 30 scientists and technologists of Indian origin (STIOs) into researcher leadership roles.

"The offer was withdrawn as he did not accept the terms and conditions and demanded unreasonable compensation," Samir Brahmachari, director general of the CSIR, told *Nature*.

Ayyadurai denies this. In a 30 October letter to Prime Minister Manmohan Singh, who is also president of the CSIR, he claims that he was sacked for sending senior CSIR scientists a report that was critical of the agency's leadership and organization. The report, published on 19 October, was authored by Ayyadurai and colleague Deepak Sardana, who joined the CSIR as a consultant in January. Ayyadurai says that the report — which was not commissioned by the CSIR — was intended to elicit feedback about the institutional barriers to technology commercialization.

"Our interaction with CSIR scientists revealed that they work in a medieval, feudal environment," says Ayyadurai. "Our report said the system required a major overhaul because innovation cannot take place in this environment."

Ayyadurai says that he came to the CSIR with a "mission" to apply his scientific and entrepreneurial experience to help his homeland, and that Brahmachari had promised him the authority, budget and resources to execute the mission. But Ayyadurai's relationship with Brahmachari soured after the report went public. Ayyadurai claims that on 23 October he was barred by the CSIR from speaking to council scientists or directors, and was denied Internet and e-mail access. His appointment offer was



Shiva Ayyadurai says he returned to India on a "mission" to help his homeland.

withdrawn on 26 October, and on 7 November he was given three days to vacate his residential accommodation provided by the CSIR. Brahmachari confirms this chain of events.

Sardana shares Ayyadurai's views. Describing their report as an in-depth study of the management of the CSIR, he wrote to science minister Prithviraj Chavan on 19 October saying that "it is not possible for me to continue working without your immediate direct intervention" because of the problems triggered by the report.

## Impervious to criticism

"I have seen many cases of vindictiveness in the CSIR, but this is the worst," says Pushpa Bhargava, founder director of the CSIR's Hyderabad-based Centre for Cellular and Molecular Biology (CCMB). Bhargava, who has also written to Singh supporting Ayyadurai, says: "Ayyadurai's report tells the truth about how the CSIR is being run today. The fact that CSIR administration is impervious to healthy and fair criticism is bound to send the wrong message not only to expatriates but also [to scientists] within the country."

"I am more worried that the incident will dampen the enthusiasm of Indian institutions to hire expatriates in the future," says Valangiman Ramamurthy, the former science secretary of the government's Department of Science and Technology, who recommended Ayyadurai's selection.

But others think that Ayyadurai's case will not set a precedent. Gangan Pratap, director of the National Institute of Science Communication and Information Resources, a CSIR institute in New Delhi, says that most returning expatriates will focus on research and teaching, rather than getting involved with policy issues, and are unlikely to face similar conflicts. Rajan Sankaranarayanan, who returned from the Institute of Genetics and Molecular and Cell Biology in Strasbourg, France, to join the CCMB in 2002, agrees that the episode should not discourage other Indians from returning, as long as they are prepared to work within the existing systems.

Brahmachari is also optimistic. "Serious scientists from top institutions around the world have shown their willingness to join the CSIR's STIO programme, not because of perks or position, but because of the intellectual environment that the CSIR offers," he says. "Right now we have requests from several top people from the United States wanting to work in the CSIR." ■

**K. S. Jayaraman**

> "I have seen many cases of vindictiveness in the CSIR, but this is the worst."

---

**Correction**

The European Research Council will appoint a leading scientist to the new post of director, which is not a direct replacement of the current secretary-general Andreu Mas-Colell as we suggested in our News Briefing item (*Nature* **461**, 1178; 2009).

# One killer virus, three key questions

As the world mobilizes against the H1N1 flu pandemic, researchers are working to answer pressing questions about the virus. **Brendan Maher** visited pathologists at the US Centers for Disease Control and Prevention who are looking at how the virus kills, and a New York laboratory that is testing how it spreads. **Declan Butler** spent time at a French biosafety level-4 facility where researchers are working out the chances that the pandemic virus will reassort with the H5N1 avian flu virus.

## 1 How does it kill?

There are five sets of eyepieces on the microscope, and Sherif Zaki is looking down one of them. Looking down the other four are members of the team he leads at the infectious-disease pathology branch at the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia. Other researchers trained in epidemiology, microbiology and electron microscopy watch a large, flat-screen monitor at the end of the narrow conference room that shows an image projected from the microscope.

Dianna Blau, an epidemiologist, reveals the source of the tissue: an 11-year-old girl who died in September, probably from H1N1 influenza. The team had already detected viral RNA in the girl's samples. But observing the presence of the H1N1 virus, especially in this tissue from deep down in the lungs, provides a more accurate and detailed diagnosis. The tissue seems messy, and is flooded with blue-stained *Staphylococcus aureus*. Zaki scans back and forth across the slide, seeking the hint of red staining that would indicate the presence of antibodies bound to the H1N1 virus. He finds it on one slide, where a red blob indicates a profusion of virus being released from a rupturing cell. They mark her as positive.

Zaki's lab holds these 'sign out' meetings every afternoon, examining autopsy and other tissues that have been sent to the CDC from doctors and medical examiners around the world to find, confirm or rule out any number of diseases. Since April, their case load has doubled thanks to H1N1. So far, the group has received samples from more than 300 suspect cases, and it has confirmed swine flu in more than a third of them.

In addition to providing a diagnosis, the group's analyses are helping to build up a picture of how the virus kills. Pathological studies can show what tissues are affected and to what extent, adding detail to the worldwide monitoring and surveillance efforts.

The profile emerging is of a distinctive virus. Although seasonal flu tends to infect just the cells high in the upper airway, H1N1 penetrates down into the terminal air sacs called alveoli. "This is not an area of the lung where you would usually see seasonal flu," Zaki says. He has seen such behaviour before, though — in the few samples of lung tissue he has examined from humans killed by the H5N1 avian flu virus. But the virus is much more prevalent in the tissues from the severe H1N1 cases he has examined — "like avian flu on steroids" as Zaki puts it.



J. GATHANY/CDC

**Sherif Zaki (front) studies tissue infected by pandemic H1N1.**

Zaki says that his observations fit well with recent research looking at the mechanism of infection. A group led by Mikhail Matrosovich at Philipps University Marburg in Germany and Ten Feizi at Imperial College London studied sialyl glycans, glycoproteins that the flu virus binds to in order to gain entry to human cells[1]. Although seasonal strains of H1N1 bind mostly to versions of the glycoproteins known as α2-6, the researchers found that the new pandemic H1N1 can also bind to a version called α2-3, which is found in greater proportion in the lower respiratory tract.

Co-infection is common with pandemic H1N1, at least in those who have died. Zaki's group has observed infection with bacteria such as *S. aureus* or *Streptococcus pneumoniae* in about a third of the fatal swine-flu cases it has examined. In the rest, the virus seems to be lethal on its own. Zaki slots onto the microscope stage a slide from a 38-year-old male who died without bacterial co-infection. This one is filled with the red staining; also the walls of the alveoli are ruptured and blood cells and fluid fill the spaces normally

reserved for gas exchange. Scar-like pink ribbons, called hyaline membranes, arc through the tissue. Pathologists call this state 'diffuse alveolar damage', and it tells them that the man had respiratory distress syndrome. "It's very difficult to treat a patient once they get to this state," Zaki says. He points out that the man was obese and had a history of hypertension and heart problems. Some 90% of the cases his team has reviewed have had some underlying medical condition.

Zaki says that they plan to publish some of their observations soon, and he has been sharing them with others at the CDC and with the public-health community at large. Although pathology can't predict what the virus will do in the future, it can help to identify those most at risk of severe disease. "We've really learned a lot with the lab pathology," says Anne Schuchat, director of the CDC's National Center for Immunization and Respiratory Diseases. "Both the prominence of the pneumonia in some infections and bacterial co-infection have important clinical implica-

> **"H1N1 looks like avian flu on steroids."**
> — Sherif Zaki

tions," she says, explaining that pathology results have led the CDC to recommend pneumococcal vaccination for people in at-risk groups.

At the end of the sign-out meeting, Zaki walks back to his office through a quiet lab; the technicians have gone home for the day and the machines used to prepare and stain slides are silent. Six fresh case studies in purple folders sit in a neat row along one of the lab benches ready for processing and examination. They contain slides or tissue embedded in paraffin: one still has lumps of formalin-soaked flesh sitting in a plastic specimen jar.

As the flu season ramps up, Zaki says, he suspects the number of flu cases will increase dramatically. Still, his lab sees only a small slice of what's happening worldwide. "We're looking at the tip of the iceberg in terms of these cases," he says. It's a grim job, reading the tales of the dead, but he adds "that information is important for the living".

**Brendan Maher**

## 2 How does it spread?

In a cramped locker room high above the streets of upper Manhattan, laboratory personnel take a visiting journalist into the emerging pathogens facility, an enhanced biosafety level 3 (BSL-3) laboratory on the New York campus of Mount Sinai School of Medicine.
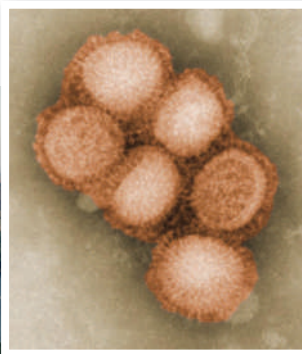
The suiting-up procedure is complicated. Rubber-soled booties must be tucked under a jumpsuit to ensure that spilled fluids don't drip into them. Purple latex gloves are layered underneath green ones; the different colours are used to ensure that rips in the outer layer are visible. Black N95 face masks pinch off the nostrils, forcing the wearer to breathe noisily through his or her mouth. The staff call them Darth Vader masks. They also wear

papery mesh scrubs and white 'bunny suit' coveralls made from the material used to insulate houses. Biosafety officer Philip Hauck quotes Dante as he prepares to go in: "Abandon all hope ye who enter here."

In another preparatory room, they add heavy, belted, battery-powered air filters with a hose that hooks into the back of a white head-covering. Once it's adjusted, air begins rushing up over the face, puffing up the suit. The positive pressure in the suit, combined with the negative air pressure in the lab, ensures that the flow of air will carry any airborne pathogens away from the body.

This is the procedure that researchers at Mount Sinai go through every day to carry out their animal studies on deadly pathogens. Since May, pandemic H1N1 has been one of those pathogens. As the pandemic was emerging, John Steel and Anice Lowen secured strains of novel H1N1 from California and the Netherlands, and made plans to include them in their work modelling the transmission of influenza in guinea pigs. BSL-3 facilities are not required to work with H1N1, but Steel and Lowen's institutional review boards wanted the work to be carried out under the most stringent conditions on site.

Many labs use ferrets to model the transmission and pathology of human viruses, but the animals can be difficult to work with. Peter Palese, the head of microbiology at Mount Sinai, came up with the idea to work with guinea pigs a few years ago after reading an article published after the 1918 influenza pandemic, in which researchers in New Mexico noted that the infection had killed off a number of laboratory guinea pigs[2]. He convinced Lowen and Steel, assistant professors in his lab, to test whether the creatures would transmit human influenza. They did. In 2006, the group showed that the guinea pigs pass the virus between them with about the same efficiency as humans[3], and the researchers still use them as a model of transmission



Biologists at the Mount Sinai School of Medicine are studying transmission of the pandemic H1N1 flu virus.

MOUNT SINAI SCHOOL OF MEDICINE; C. S. GOLDSMITH & A. BALISH/CDC

even though their animals, purchased from a lab supplier, do not show symptoms or die from the disease.

The first thing the researchers did with the samples of pandemic H1N1 was to compare its transmissibility with that of seasonal flu. They squirted about 10,000 infectious H1N1 particles up the noses of four guinea pigs and, a day later, placed these animals in cages next to uninfected animals, separated by wire-mesh walls that allow respiratory droplets to pass through. Within five days all the exposed animals tested positive for the virus. The team's paper, published last month in the *Journal of Virology,* suggests that the H1N1 swine flu virus transmits just as efficiently as seasonal flu[4]. This is contrary to some work on transmission done at the start of the pandemic[5] but mirrors real-world data showing that H1N1 spreads rapidly. "The pandemic sort of scooped us," says Lowen. Palese's group also found that previous exposure to the H1N1 and H3N2 subtypes of seasonal influenza limits the ability of exposed animals to become infected with the pandemic H1N1, which supports the idea that a seasonal flu infection or vaccine might offer some cross-protection against swine flu.

Inside the facility are the huge steel ferret cages that Steel and Lowen have adapted for their guinea-pig work so that two animals can fit inside each. They are now planning new experiments, hoping to find the sequences within the genome that enable the virus to transmit so well between humans compared with other swine-adapted influenza strains. They may create reassortments of the virus and then test how well they are transmitted by the guinea pigs. The results might aid in surveillance efforts aimed at identifying other strains from animals that could make the leap to humans.

Everything that goes into the BSL-3 lab must be decontaminated before coming out. This involves spraying down the bodysuit with ethanol, carefully peeling it off and then leaving the facility by way of a shower room. A pen and notepad won't survive the caustic chemical dunk by the exit, so the pages are scanned and put onto a CD that can be sprayed down.

Outside again on the Manhattan streets, H1N1 could be encountered in any cough or sneeze. Without all the protective gear, one feels rather vulnerable.
**Brendan Maher**

# 3 What could it turn into?

A deadly line-up of viruses is locked up in the computer-controlled safes at the Jean Mérieux/INSERM biosecurity level four (BSL-4) facility in Lyon, France, including Ebola, Nipah, Lassa, Hendra and Marburg. And in the next few weeks, scientists working there are planning to manufacture a new resident. They hope to test whether the highly transmissible pandemic H1N1 virus could reassort with its deadlier cousin, the H5N1 avian flu, to make a virus with the worst properties of both.

Classed as a national high-security facility, the laboratory is a three-storey shoebox made from armoured glass, perched above a biology research centre of INSERM, the national biomedical agency. The edifice is built to withstand earthquakes, bullets and explosives. It is also smack in the city centre, its entrance just metres from mothers pushing prams along the pavement on avenue Tony Garnier.

The lab itself is surrounded by empty corridors and a spartan decor of steel staircases, trusses, raised walkways and ventilation pipes. The air hums from the air-ultra-filtration and other support machinery that take up entire floors above and below. Under negative air pressure to stop viruses escaping in the event of a breach, the lab is split into three zones, with the airstream flowing towards the 'hottest' exposure risk zone, the animal house at the back.

When the current H1N1 pandemic began in April, priority lab time here was allocated to Bruno Lina, a virologist and flu researcher at the CNRS, France's basic-research agency, who works at the University of Claude Bernard Lyon-1. Pandemic H1N1 influenza is not itself a BSL-4 agent — a BSL-2 facility is adequate. But in France, the health ministry classes viral reassortment experiments of the sort Lina is performing as requiring BSL-4 precautions.

The ability to predict which reassortments might take place, and what type of flu viruses might result, could be key to predicting the behaviour of this and future pandemics. The eight genes in the influenza virus's segmented genome are easily swapped between strains. But to be viable, new gene combinations must also be able to work together to package themselves into a virus particle. Viruses within a subtype tend to reassort with one another more easily and generate more viable reassortants than do those from different subtypes.

Over the summer, Lina's team has been using the BSL-4 facility to investigate the likelihood that pandemic H1N1 will acquire resistance to the front-line antiviral drug oseltamivir (Tamiflu) through reassortment, and how easily these reassortants might spread. Resistance can emerge by spontaneous mutation, but given that seasonal H1N1 is already resistant to the drug and spreads easily, reassortment is perhaps the most likely way that pandemic H1N1 will acquire resistance — especially as seasonal H1NI and pandemic H1N1 are the same subtype. Since the start of the pandemic, Tamiflu-resistant strains have sporadically appeared in several countries but none has yet gained a foothold. That



P. LATRON/INSERM

The biosafety level-4 facility in Lyon, France.

Suited up: extensive precautionary measures have to be followed in biosafety level-4 labs.

they haven't arisen more often or spread more easily may be because there is little seasonal H1N1 circulating, as pandemic H1N1 is outcompeting it — a large number of co-infections are needed for transmissible reassortants to arise.

In his work, Lina co-cultured the two H1N1 viruses in a cell line. He is now testing how pathogenic the reassortants are in mice, before using ferrets to test their capacity to spread. The biggest hazard that the scientists face in the lab is a bite or a scratch from an infected animal, so all manipulations are done on animals that have been anaesthetized remotely.

Lina is also preparing to submit a protocol to the facility's scientific board seeking the green light to try to reassort pandemic H1N1 and the H5N1 avian flu virus. "It's controversial research, but it is basic science that needs to be done," says Lina. H5N1 has killed more than half of the people it has infected since it resurged in 2003, but has rarely spread from one person to another. H1N1, on the other hand, seems to be as transmissible as seasonal flu, but mild in most people, if severe in some. The aim of Lina's proposal is to find out the probability of a reassortant arising that combines the lethality of H5N1 with the transmissibility of H1N1.

In particular, Lina is searching for putative molecular controls of the virus's segmented genome that may determine why some reassortants can package the virus and others can't. "We don't know which regions in H5N1 are responsible for that control," says Lina. That could have benefits for disease surveillance, he says. If researchers know the key genetic regions that facilitate reassortment, surveillance efforts could watch out for H5N1 or H1N1 viruses with changes in those regions, ones that might be on the verge of dangerous reassortments.

Lina will use reverse genetics to generate a soup of reassortants, test whether any are viable, and if they are, assess their virulence and transmissibility. Because H1N1 and H5N1 are different subtypes, Lina does not expect them to swap genes easily. In 2005, he tried to reassort H5N1 with seasonal H1N1 and H3N2 viruses, without success. "After a year we only had three reassortants, and none was fit," recalls Lina, "they just don't reassort well."

The experiments that Lina's team is carrying out in the BSL-4 lab aren't technically novel or difficult in themselves.

> **"It's controversial research, but it is basic science that needs to be done."**
> — Bruno Lina

But the encumbrance of the safety procedures for getting in and out of the high-security lab make even the most straightforward procedures complicated, says Lina. Looking through the bulletproof window, one can see why. As the researchers move around the lab, they connect and disconnect their blindingly white spacesuits from some 60 yellow air hoses dangling from the ceiling. The hands of the researchers snipping away at tissue samples are wrapped in multiple layers of gloves to avoid accidental pricks. "It's like trying to do surgery wearing boxing gloves," says Hervé Raoul, director of the facility.

Before researchers can even begin to work in the lab, they must pass a three-week training course. It's only after a further 200 hours of practice that they are given any real autonomy in the lab, and they are never allowed inside alone. Users are limited to working a maximum of one four-hour shift a day, so that they are less likely to make mistakes. They are also encouraged to share any concerns about the behaviour of colleagues, or whether they themselves are going through a difficult patch, which might cause distraction and accidents.

Before Lina's experiments get approved, the facility's external scientific board will need to be convinced of the public-health justification, that the science is top-notch and that the experiments can be done safely, says Raoul. It will also need to be approved by government regulatory agencies. If all goes well, he could have his authorization in weeks, Lina says.

"Compared to the big guns in flu research, I'm a little guy," says Lina. "But having access on my doorstep to one of the rare BSL-4 facilities in the world with an animal house is a big advantage." ∎

**Declan Butler**

1. Childs, R. A. *et al. Nature Biotechnol.* **27,** 797–799 (2009).
2. Lamb, F. H. & Brannin, E. B. *J. Am. Med. Assoc.* **72,** 1056–1062 (1919).
3. Lowen, A. C., Mubareka, S., Tumpey, T. M., García-Sastre, A. & Palese, P. *Proc. Natl Acad. Sci. USA* **103,** 9988–9992 (2006).
4. Steel, J. *et al. J. Virol.* doi:10.1128/JVI.01732-09 (2009).
5. Maines, T. R. *et al. Science* **325,** 484–487 (2009).

**See Editorial, page 137, and News, page 146.**

# CORRESPONDENCE

## Weapons: the need to replace ageing and deteriorating stock

SIR — Your Editorial on resistance to non-proliferation of nuclear weapons is, in my view, a good example of selective presentation of data (*Nature* **461,** 11; 2009). As first director of the US Defense Threat Reduction Agency, from 1998 to 2001, I consider a few corrections are in order.

Although the predicted lifetimes of the plutonium pits of US weapons are 40–80 years, the thousands of non-nuclear components in the pit's environment are less stable. Radiation damage, loss of plasticizers and binders, and other changes mean that they could eventually compromise the performance of highly optimized cold-war weapons, designed to minimize weight and use of plutonium and uranium.

Replacing these parts with exact equivalents is often not possible because the materials and production processes no longer exist or because their production is precluded by changes in environmental and safety regulations. Confidence in performance and the willingness of politicians to accept expert judgement about it are likely to suffer.

The proposed Reliable Replacement Warhead (RRW) was designed using nuclear systems that were more robust and had higher margins against failure, thus relaxing the stress on new non-nuclear systems intended for replacement and future production. As a bonus, the safety and surety features of the weapon were improved. It is not a stalking horse for nuclear testing but would increase military and congressional confidence in weapons performance; this assurance is vital as falling numbers render each weapon more important for deterrence.

Unlike the unlamented Robust Nuclear Earth Penetrator, which failed tests of both utility and credibility, the RRW does not represent a lowering of thresholds against nuclear use, nor an opening to worrying policies of pre-emption.

You should not chide the United States for desiring to do what the British, French, Russian and Chinese governments are also doing or planning to do — namely, modernizing their stockpiles with new weapons systems.
**Jay Davis** Buena Vista Avenue, Livermore, California 94550, USA
e-mail: jcd42@sbcglobal.net

## Weapons: existing stockpile can be safely maintained

SIR — You are correct to state that the Reliable Replacement Warhead (RRW) programme would do nothing to improve the reliability of the US nuclear arsenal (*Nature* **461,** 11; 2009). It is possible that the "powerful figures within Obama's own administration" are being advised by people associated with the weapons laboratories who may have an interest in derailing a ratification of the Comprehensive Test Ban Treaty.

Existing nuclear weapons are already very reliable and their safety features are adequate, as a series of reports by R. E. Kidder of Lawrence Livermore National Laboratory affirmed. (For further details, see go.nature.com/yCmG4W.) As US weapons are comfortably tolerant of small variations that may occur in materials or the manufacturing process, they can be remanufactured without explosive proof-testing.

Those in favour of the RRW claim that it would not need nuclear testing. This is because the RRW would be composed of an existing primary stage, which is where many of the uncertainties of nuclear weapons design reside, and probably also an existing secondary stage. However, very few people outside the weapons-design community would trust deployment of a weapon that has not been tested.

As a result, the pressure to test any RRW could well derail the ratification of the test ban treaty. Failing to ratify this treaty (not to mention resuming nuclear testing) would do great harm to US national interests as well as those of the world at large.

Others argue that the RRW programme is needed to maintain expertise in weapons physics. This can be done without the RRW. There is no need for it or any continued nuclear proof-testing.
**Gerald E. Marsh** East View Park, Chicago, Illinois 60615, USA
e-mail: gemarsh@uchicago.edu

## Many types of action are required to tackle climate change

SIR — Sanjay Khanna's survey of climate-campaign activities (*Nature* **461,** 1058–1059; 2009) implies that the arts and advertising ought to be helping to bring into being a "worldwide consensus for action". But the point is not to engineer a global consensus for action, as though the 'action' that is being sought is somehow self-evident, unambiguous and uncontested. Palpably, it is not.

Instead, the urgency is to articulate the many types of action — individual or collective, selfish or altruistic, conservative or radical — that can be justified by the prospect of climate change, and to understand why, in a plural and contradictory world, these actions may differ. And then to accommodate them.

To paraphrase one of the six messages that came out of the Copenhagen climate congress in March, it is not inaction that is inexcusable. What is inexcusable is to pretend that there is just one message, one voice, one number to be communicated, that there is just one action agenda that (the science of) climate change demands from us, and that the arts and advertising can deliver it.
**Mike Hulme** School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK
e-mail: m.hulme@uea.ac.uk

## Decarbonization figures for India and China unconvincing

SIR — In their Opinion articles, Rajendra Pachauri (*Nature* **461,** 1054; 2009) and Jiahua Pan (*Nature* **461,** 1055; 2009) include figures that show business-as-usual perspectives on India's and China's emissions growth in coming decades. Both figures indicate that the two countries' economies have already made the transition to high rates of decarbonization. This would put them in a very strong position in international climate negotiations, but I find the figures unconvincing.

India's five business-as-usual projections include different assumptions of annual rates of decarbonization, from 1.0% up to 3.3%. Four of these greatly exceed the 1987–2006 average of 1.1%. The single business-as-usual projection for China suggests an annual rate of decarbonization of 6.5% per year to 2030, which is almost three times the 1987–2006 average. China's emissions grew by 12.2% per year from 2000 to 2007, and under business-as-usual are now expected to grow by only 2.5% per year to 2030.

If India and China have indeed already implemented policies that will decarbonize their economies by 3% per year and more, then it would be very good news indeed, as global rates of about 5% (or more) per year would be necessary to achieve an 80% emissions reduction below 1990 levels by 2050, assuming modest economic growth. But if India and China are overoptimistic about future rates of business-as-usual decarbonization, then the challenge of stabilizing concentrations of carbon

**"About 70% of women holding degrees in science, engineering or technology do not currently work in these fields."** Asha Gopinathan, page 165

dioxide will be much greater.

Either way, projections of business-as-usual decarbonization from any country that are at rates three times higher than recent historical averages should be greeted with appropriate scepticism.

**Roger A. Pielke Jr.** Center for Science and Technology Policy Research, UCB 488 University of Colorado, Boulder, Colorado 80309-0488, USA
e-mail: pielke@colorado.edu

## No special cases in efforts to stop immigration fraud

SIR — As a scientist working in a foreign laboratory, I can understand Colin Stoneking's frustration at the UK visa system for students (*Nature* **461,** 1053; 2009). However, there is a reason for the unwieldy processes he complains about.

They have been developed to combat the problem of colleges advertising questionable courses to overseas students. In some cases, these are merely fronts to allow economic migrants to circumvent UK border controls, or to defraud overseas students of their tuition fees. A report by the UK Parliament's Home Affairs Committee found that of roughly 4,000 establishments catering to overseas students, 2,200 had not been licensed to sponsor students under the new immigration points system (see go.nature.com/9LXmvd). The implication was that some of them may be, to quote immigration minister Phil Woolas, "dodgy".

The progress of science has always been, and will continue to be, aided by the free exchange of ideas and people between nations. But in the present economic and political climate, we should not expect to be treated as a special case.

**Iain Scott** National Heart, Lung and Blood Institute, 9000 Rockville Pike, Bethesda, Maryland 20892, USA
e-mail: scotti@nhlbi.nih.gov
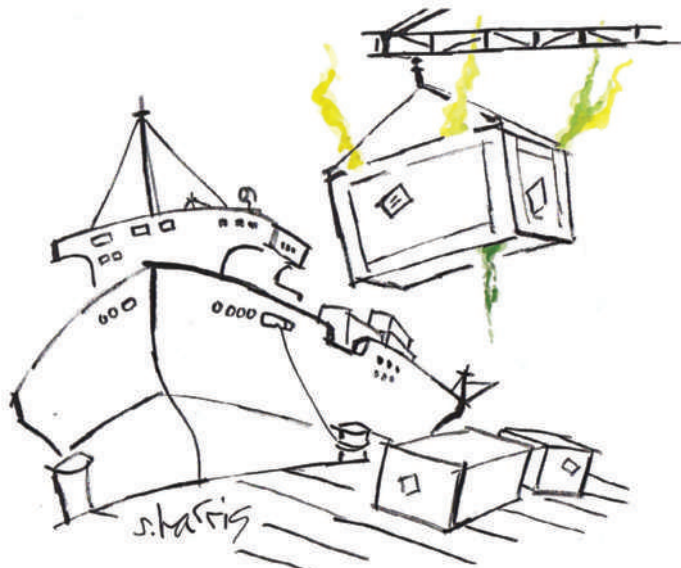
## Emissions affected by trade among developing countries

SIR — As the climate-change conference in Copenhagen approaches, attention is focused on the dialogue between developed and developing countries (see www.nature.com/roadtocopenhagen). But the impact of the growing links between developing countries should not be overlooked.

China is investing heavily in African countries, focusing mainly on improving their infrastructure, as shown by the *China Statistical Yearbook 2009* (National Bureau of Statistics of China, 2009). From 1991–2007, imports to China from Africa — chiefly raw materials and emission-intensive products — grew by a factor of 45, compared with just a sixfold increase in imports from the European Union.

Half of China's recent increase in carbon emissions has been driven by its production of goods for export, 60% of which went to wealthy Organisation for Economic Cooperation and Development (OECD) states (D. Guan *et al. Geophys. Res. Lett.* **36,** L04709; 2009). However, China's imports from Africa are responsible for rapidly increasing African carbon emissions from less than 400,000 tonnes to more than 40 million tonnes between 1991 and 2008 (*China Statistical Yearbook 2009* and G. P. Peters *et al. Environ. Sci. Technol* **42,** 5; 2008). These already account for 5% of Africa's total emissions.

Addressing global greenhouse-gas emissions requires action from developed countries, but also cooperation by developing countries. The more prosperous developing nations have a critical role: by linking resource providers and consumers they can funnel the spillovers of low-carbon technologies to the less developed. Success depends on large-scale financial and technological flow, initiated by OECD countries, with an effective mechanism for sustaining a strong increase in resources and in technology.

**Dabo Guan, David M. Reiner** Electricity Policy Research Group, Judge Business School, University of Cambridge, Cambridge CB2 1AG, UK
e-mail: dg346@cam.ac.uk

## Phosphorus decline could be good for water supplies

SIR — There is an upside to the potential future phosphorus shortage highlighted in your News Feature 'The disappearing nutrient' (*Nature* **461,** 716–718; 2009). Long before a global phosphorus crisis occurs, declining supplies may present a windfall for water quality and sustainable agriculture.

As you point out, declining phosphorus supplies will encourage efficient recycling of phosphorus-rich livestock manure. The full potential of livestock manure fertilizers is not being achieved, owing to unfavourable economics and nutrient ratios that are incongruent with crop needs. The phosphorus-to-nitrogen ratio of livestock manure (roughly 1:3) greatly exceeds the phosphorus-to-nitrogen ratio required by common cereal crops such as maize and rice (roughly 1:6). And on multi-operation farms, manure is free.

As a result, manure fertilizer applications often exceed crop phosphorus requirements to meet crop nitrogen demands. The excess phosphorus can be transported to surface waters, fuelling eutrophication of freshwater lakes and low-salinity estuaries such as Lake Victoria and the Stockholm archipelago.

Declining phosphorus supplies coupled with increasing phosphorus demand will encourage sustainable agricultural practices that include more effective manure-fertilizer management. For example, livestock manure can be supplemented with synthetic nitrogen fertilizer to narrow the gap between manure and crop nutrient ratios. Better matching of available-nutrient supply with crop-nutrient demand promises to extend global phosphorus supplies and improve water quality throughout the world.

**Michael J. Castellano** 116 ASI Building, Department of Crop and Soil Science, The Pennsylvania State University, University Park, Pennsylvania 16802, USA
e-mail: mjc471@psu.edu

Contributions to this page may be submitted to correspondence@nature.com. We also welcome comments at Nautilus (http://blogs.nature.com/nautilus).

# OPINION

# Universities need a new social contract

To reconcile solution-driven research and blue-skies thinking, academic institutions urgently need innovative collaborations and new funding models, says **Indira V. Samarasekera**.

Over the past year, academic leaders from around the world have met to contemplate the future of higher education and university research, against the backdrop of global financial upheaval. As president of the University of Alberta in Edmonton, Canada, I have participated in some of these international roundtables. My conclusion? It is time to construct a new social contract between research universities and their public and private partners: one that both promotes the pursuit of basic research and encourages solution-driven work. We, the academic leaders and universities, should embrace this new relationship, establishing a funding mechanism to fit. We should devise new ways to measure success, and actively copy the organizations that work best. If we can do all that, we stand a better chance of solving the world's problems — now and in future.

Public and private universities worldwide face a bleak financial future. The value of endowments has plummeted — by as much as 30% in some institutions over the past year. Government support has been slashed — by up to 20% in the United States and up to 8% in Canada. Philanthropic support is harder to come by, and aspiring students can barely afford current fees, let alone increases in debt.

Yet nations are counting on the talent of graduates and on the discoveries of university researchers to restore and advance the global economy. As testimony to this faith, investments in research and development feature in stimulus packages — including US$16 billion in the United States and Can$2 billion in Canada. Many of these investments are targeted to support solution-driven research in specific areas. The US stimulus package includes funding for advanced energy research and climate-change research. The Canadian package funds ready-to-be-built infrastructure projects at colleges and universities. In addition, Can$200 million over seven years has been committed to the Canada Excellence Research Chairs, a new federal programme to attract top academics to build world-class centres. These hubs will focus on areas identified as strategically important to Canada's long-term economic plan: environment, natural resources, life and



health sciences and information technology.

This push towards more solution-driven research funding, which pre-dates the recession, is a source of growing concern for many academic researchers, and for good reason. They are worried about the potential devaluation of basic-science research and arts scholarship, which have led to profound advances in human knowledge and to major commercial successes. Such 'blue-skies' research was, until recently, considered the mainstay of universities and a crucial part of the education of undergraduate and graduate students, and it must remain so.

But converging forces — the expansion of globalization, the increasing ease of communica-

tion and the trend of 'open innovation', whereby companies promote research outside their own buildings — are reshaping how public universities work, and to what end. Understandably, the supporting populace, governments, industries, philanthropic organizations and social agencies are calling for researchers to focus on seeking solutions to specific challenges.

Some researchers have responded enthusiastically. But academic thinking and funding mechanisms have not kept pace with the dual imperatives of blue-skies research and solution-driven research. It is time to bridge this gap.

## Slow going

The most urgent problems demanding scientific and technological research attention today are global — from international security to energy, environmental sustainability and economic recovery. To be fast and effective, we must stimulate and support interdisciplinary, inter-profession and inter-sector approaches, funded internationally.

What obstacles stand in the way? Inertia and

---

**SUMMARY**

- Answers to big global problems are being lost to structural inertia
- Interdisciplinary, inter-institutional, international projects need support
- The world's government funding leaders must design a fix together

C. LAW

inflexibility. Universities joke about operating in glacial time frames, but we aren't alone. Funders also contribute to delays, requiring repeated rounds of negotiating, reworking of proposals and multiple levels of review. This needs to change. A new social contract would involve establishing a review, approval and funding process for solution-driven research that sits parallel to processes for basic research, and has the nimbleness and urgency desired in the outcomes, without compromising the cautions that validate and protect the quality of the work.

At present, researchers, faculties and universities fight for research funding, ranking and reputation. We must redirect this competitive energy by using funding models that encourage collaboration. If we wish to speed the passage from problem to solution, we need to play with, rather than against, each other.

Universities have a bigger part to play in today's world. With major companies doing less of their own basic research, universities are increasingly becoming repositories of the specialized environments, state-of-the-art equipment and talent needed for blue-skies and solution-driven research. Companies and governments increasingly depend on them for technology development and innovation.

Universities and companies are recognizing that innovation is a contact sport. We need to encourage and support collaboration between universities and the private sector. Efforts to bring university researchers into more frequent contact with industry, to better identify profitable challenges, need to be strengthened — informally and formally — through partnership agreements.

Already a few organizations embody the tenets of the new social contract as I've outlined them. The Worldwide Universities Network, Academic Consortium 21 and the International Alliance of Research Universities, for example, encourage collaborations across national and disciplinary boundaries, assist in developing online research networks and consortia, and facilitate faculty and student exchanges. The long-term success of these organizations will depend on each member institution's willingness to provide support (personnel and money) to participating faculty members.

A good example of a nimble funding body is the Bill & Melinda Gates Foundation's Grand Challenges in Global Health programme, which focuses on finding quick and effective solutions to identified problems and cuts through some of the red tape that typically causes funding delays. Application forms are short; initial US$100,000 grants are awarded within three months of submission; and those projects that go on to produce results can apply to receive

up to $1 million more. The programme does not discriminate between applications from students or well-established researchers and it encourages proposals that take unconventional, interdisciplinary approaches.

The European Research Council, established in 2006, is another cross-border organization that promotes 'frontier' science and encourages applications from early career researchers as well as from established scholars. But the council has a fairly traditional structure, with a focus on curiosity-driven projects developed by principal investigators. There are no barriers to developing international or interdisciplinary research teams, but neither are there special inducements to do so. This seems to be a lost opportunity, especially in an organization that arose, in part, from a desire to unite research efforts across European countries.

## 'Collaboratories'

Looking to public–private partnerships, we can see other promising developments that encourage both basic and solution-driven research. In Canada, Mike Lazaridis, co-chief executive of Research in Motion in Waterloo, Ontario, and inventor of the BlackBerry communication tool, founded the Perimeter Institute of Theoretical Physics adjacent to the University of Waterloo. He has made a commitment to support blue-skies thinking on the nature of the Universe with investments of more than Can$150 million, even though the discoveries are not likely to benefit his business for the foreseeable future. At the same time, Microsoft, Intel, Xerox, Tata, BP and BT have labs or researchers in universities to foster co-discovery and innovation.

The National Institute for Nanotechnology (NINT) located at the University of Alberta presents yet another model: a 'collaboratory' of university, government and industry interests. The researchers at NINT undertake basic research on nano-materials and devices but also work with companies such as Hitachi to develop the next generation of microscopes. IBM is similarly implementing collaboratories on a global scale.

Managing the balance between the research and publication needs of faculty members and graduate students with the needs of industry to patent and develop products is never easy, but the benefits outweigh the risks. Resources are leveraged, students are supported with scholarships and internships, researchers gain insight to help set research priorities that best address real-world challenges and companies get a first shot at applying

research breakthroughs to new products.

These pockets of promise are emerging organically and, in some cases, rather haphazardly, without an over-arching multinational infrastructure or drive. Without common vision, guidelines and metrics for measuring progress, it is difficult to determine which strategies are actually succeeding and to build on them.

There are two ways we can respond to this situation. We can continue to muddle forward and hope for the best. If so, we risk long delays — or failure — in achieving the results that the world needs now. Researchers eager to create interdisciplinary, international collaborations will continue to find it difficult, sometimes impossible, to find the funding they need. The alternative is to be proactive: to analyse the success of emerging models and copy strategies that work.

I suggest we begin by bringing together leaders from major government funding agencies across the world — particularly from North America, Europe, India and China. Together they need to define a funding model designed to award interdisciplinary, inter-institutional and international projects. They would then need to determine what percentage of their funding budgets could be reallocated to these kinds of projects. They must also develop systems for efficient submission and evaluation of proposals and for distribution of funds. The format for these discussions could be guided by lessons learned during the establishment of the European Research Council and the ongoing Bologna Process, which is working to standardize degree requirements and quality controls in universities across Europe.

> **"If we don't forge a new social contract now, we will bequeath to the next generation an even more ramshackle system."**

Some may say that a time of economic uncertainty is not a good moment to instigate change, but I believe there will never be a better time. I think of the research economies to be gained through collaboration and integration across boundaries. I see the speed with which global problems could find solutions, once old obstacles are removed.

If society doesn't seize the opportunity to forge the new social contract now, while economic pressures strengthen our will to change, we will bequeath to the next generation of researchers an even more ramshackle system than we are coping with now. And many discoveries that might answer our big global problems will be lost to inertia and inaction. ■

**Indira V. Samarasekera** is president and vice-chancellor of the University of Alberta in Edmonton, Alberta, Canada.
e-mail: indira.samarasekera@ualberta.ca

# OPINION

# Global Darwin: Revolutionary road

In China, under the threat of Western imperialism, interpretations of Darwin's ideas paved the way for Marx, Lenin and Mao, argues **James Pusey** in the third in our series on reactions to evolutionary theory.

Charles Darwin's banner was first unfurled in China during the Reform Movement of 1895–98, in response to China's defeat in the Sino–Japanese War. This had been the most crushing moment in what the Chinese call their century of humiliation, during which the Manchu Qing Dynasty barely survived five great rebellions, and lost four wars against foreign imperialists: Britain, Britain and France, France, and — most galling of all — Japan. This last defeat was the most frightening, not because the Chinese feared 'puny Japan', as they often called it, but because they feared that the European powers, emboldened by this demonstration of weakness, would "carve up the Chinese melon" into colonies.

The watchword of the reform movement was '*bianfa*', meaning 'change our institutions'. But the very word 'change' was anathema to the conservative officialdom of China. So reformers turned to Darwin as a foreign authority on change, presenting him not first and foremost as a natural scientist who had discovered an amazing fact of life, but as a political scientist who had discovered a cosmic imperative for change.

Meanwhile, the Europeans waved Darwin's banner to justify imperialism. Dubbing themselves 'the fit', they declared their right to rule the 'unfit'. And some Chinese accepted this argument. Liang Qichao, one of the leading reformers, said in 1898: "If a country can strengthen itself and make itself one of the fittest, then, even if it annihilates the unfit and the weak, it can still not be said to be immoral. Why? Because it is a law of evolution."

The reformers had to find hope in *On the Origin of Species*. And they did, but their most optimistic interpretations were based on a handful of mistranslations, themselves based on a series of misunderstandings. (Westerners leapt to these misunderstandings as well — without the benefit of mistranslations.)

Chinese readings of Darwin inspired two groups — reformers and revolutionaries — to attempt to change their society through different means. Ultimately, after the failure of both groups and an erosion of traditional philosophies, Chinese Darwinian thinking prepared the nation for the rhetoric of Karl Marx, Vladimir Ilyich Lenin, and Mao Zedong.

Darwin**200**

The man who introduced Darwinian evolution to the reformers of 1895 was Yan Fu. Yan had graduated from the naval academy at Fuzhou, and was sent to England in 1877 for further study of the naval arts, which patriots hoped would one day help drive European imperialists out of the China Sea. But in England, Yan discovered political philosophy, which he came to think of as the true secret of Britain's 'fitness'. He returned to China in 1879 with a bundle of books, by Adam Smith, John Stuart Mill, Darwin and others, that he thought could rescue China from extinction. He meant to translate them, but did not publish anything for about 15 years, when goaded into action by the insult and injury of the war with Japan.

In 1895 Yan published his first essay, *Whence Strength?*, soon followed by a brilliant, periphrastic translation of Thomas Huxley's *Evolution and Ethics*, which Yan wrote in such elegant classical Chinese that even conservatives respected the text. The *Origin* itself was too long and too difficult for Yan to tackle. But Yan's translations were enough to introduce to China the basic ideas of evolution and, more importantly, the handful of Darwinian slogans that were taken up by social Darwinists around the world.

Subtle errors of translation, however, went with them: "natural selection" came out as "natural elimination"; the "survival of the fittest" became the "superior survive and the inferior are defeated"; and, causing the most confusion of all, "evolution" became *jinhua lun*, "the theory of progressive change". Strictly speaking, Darwin did not prove that evolution led to progress; to this day, that mistranslation makes it hard to discuss evolution in Chinese. As Confucius said: "If terms are not correct, discourse is difficult."

Just before Darwin's ideas reached China, the scholar and bureaucrat Kang Youwei argued that Confucius had delineated three stages of world progress: chaos, ascending peace and great peace. A mixture of these ideas soon spawned a plethora of 'stage theories' of history, all of which claimed to outline the natural and inevitable development of all 'fit races'. This seemingly benign idea — that cosmic forces

> **"Chinese readings of Darwin inspired two groups — reformers and revolutionaries."**

or natural laws were perfecting mankind and human society — eventually led to racist or class philosophies that killed people.

**Natural law**

At the start of the reform movement, the promise of Darwinian progress (which was not really Darwinian) seemed to hold the key to China's salvation. China was in the monarchy stage and should hence move on to the constitutional monarchy stage. The fittest nation on Earth, Great Britain, had shown the way.

Yan wanted democracy for China — even anarchic democracy, without presidential rule. In *Whence Strength?* his call for reform was revolutionary: "Establish a parliament at the capital and let each province and county elect its own officials." But 'Darwin' held him back from real revolution. Yan believed that step-by-step progress was a fixed natural law, so stages had to be taken in order. America had skipped constitutional monarchy and gone straight to democracy, but a resulting class war, he felt, would be their undoing. "Should we, then, now throw away all loyalty to our ruler?" he asked in his essay. "We most certainly should not! Because the time has not arrived. … Our people are not yet ready to rule themselves." (An argument that Chinese governments have used ever since to postpone democracy.)

Sun Yat-sen, later dubiously dubbed the father or George Washington of his country, was also a professed Darwinian, and an advocate of democracy. But Sun was as convinced that Darwinism was for revolution as Yan, Kang and Liang were convinced that it was for reform. One of Sun's followers, Zou Rong, put it most succinctly: "Revolution is a law of Evolution."

Taking advantage of the war against Japan, Sun and his would-be revolutionaries put their philosophy into action in 1895, hiring an 'army' from a secret society in Hong Kong in an attempt to capture the city of Guangzhou and trigger a revolution. It was an almost farcical failure. They arrived in Guangzhou by ferry, but their weapons were on the wrong boat, leaving them unarmed for their grand revolution. The police easily quashed them, although Sun managed to escape, eventually to Japan.

A few years later, the reformers had only a

G. LAM

little more luck. They won the ear of the young Guangxu Emperor and established a constitutional monarchy — on paper — in the summer of 1898. But the Emperor's aunt, the Empress Dowager, crushed the reform movement, beheaded the six leaders she could catch and put the Emperor under lifelong house arrest. Yan was somehow left alone. Others, led by Kang and Liang, took refuge in Japan.

The two self-professed Darwinian camps had much in common. Both believed in stage theories of history. Both were for democracy — but not yet. Confucian philosophy led them to believe that the fit were those who made themselves fit, and Daoist thinking made it easy to believe in a natural path that the fit could follow to survive. Both believed, at once, in determinism and 'determinationism' — a potent if illogical mix that at once met impatient patriots' demand for action and promised victory.

Sadly, both camps also accepted the pervasive Western view that Darwin had proven races unequal — that one race was 'fitter' and therefore better than another. The reformers had originally done so to disassociate themselves from those who had fallen prey to the imperialists, such as the Africans and Indians. But in their exile in Japan, reformers and revolutionaries alike turned angrily on the Manchus as scapegoats, labelling them evolutionary low life, whose 'unnatural' conquest of the Han Chinese was responsible for China's peril.

There were also crucial differences between the camps. The reformers, despite everything, remained loyal to the Guangxu Emperor. They were convinced that the stage of constitutional monarchy could not be skipped, and

they were against civil war. The revolutionaries believed that the Qing dynasty needed to be overthrown, that China could '*lie deng*' (leap over stages) to catch up to the West and that civil war was an indispensable precondition of China's evolution or progress. For a decade the two groups debated the reform or revolution question in Chinese journals smuggled back to China from Japan — with both sides wildly waving Darwin's banner.

## Enter the Marxists

In the end, the debate between reformers and revolutionaries was settled by a nearly accidental success. On a truly dark and stormy night in October 1911, Sun's followers pulled off an uprising in Wuchang. The dynasty soon fell, with Yuan Shikai, the leading 'loyalist' general, bought off with a gift of the presidency. But Yuan killed the Republic by trying to crown himself Emperor. Yuan's generals baulked. Yuan died. China fell to pieces, ruled by warlords. The debate over reform and revolution revived. Eventually the New Culture Movement arose, with Darwinian reruns, as reformers gave up on politics, embracing instead cultural reform — until 1919, when the Western powers betrayed China when they signed the Treaty of Versailles at the end of the First World War, granting Germany's 'possessions' in China to Japan.

In all of this, "politics", as Mao would later say, were "in command". Few Chinese seemed shocked by the fact of evolution, or indeed overly interested in it. Unlike Europeans, few perceived, at first, any threat to their traditional philosophies or religions. But in the decade that followed the failure of the Reform Movement,

Chinese philosophies — Confucianism, Daoism and Buddhism — did come under attack, as pacifistic doctrines that were unfit because they had rendered China unfit to survive. And so, both philosophically and politically, reformers and revolutionaries together created a naturally abhorrent vacuum.

Many tried to fill it: Sun, Jiang Jieshi (Chiang Kaishek) and, finally, the small group of intellectuals who, in indignation at the betrayal at Versailles, found in Marxism what seemed to them the fittest faith on Earth to help China to survive.

This was not, of course, all Darwin's doing, but Darwin was involved in it all. To believe in Marxism, one had to believe in inexorable forces pushing mankind, or at least the elect, to inevitable progress, through set stages (which could, however, be skipped). One had to believe that history was a violent, hereditary class struggle (almost a 'racial' struggle); that the individual must be severely subordinated to the group; that an enlightened group must lead the people for their own good; that the people must not be humane to their enemies; that the forces of history assured victory to those who were right and who struggled.

Who taught Chinese these things? Marx? Mao? No. Darwin. ∎

**James Pusey** is a professor of Chinese Studies at Bucknell University, 701 Moore Avenue, Lewisburg, Pennsylvania 17837, USA. He is the author of *China and Charles Darwin* (1983) and *Lu Xun and Evolution* (1998).
e-mail: pusey@bucknell.edu

# BOOKS & ARTS

# Fun with nuclear reactors

Two books reveal the spirit of adventure behind the history of nuclear technology, finds **William J. Nuttall.**

With a resurgence of interest in the peaceful use of nuclear power, and growing concern about the proliferation of nuclear weapons, the time is right to look back at nuclear science's turbulent history. Two recent books do so in different ways: historian of science Amir Aczel writes well for a mass readership, and nuclear scientist James Mahaffey offers extra insights for those more familiar with the story.

Aczel's *Uranium Wars* centres on developments in the early twentieth century, from the understanding of the atom to the bombing of Hiroshima and Nagasaki. Focusing on a small cast, including physicists Lise Meitner, Enrico Fermi and Werner Heisenberg, the book highlights the race to investigate nuclear phenomena. Human dramas, such as the anti-semitism and sexism experienced by Meitner, run alongside.

Although his populist treatment is necessarily superficial, Aczel uses the latest historical evidence and gives readers a trail of citations, albeit from a narrow range of sources, such as Ruth Lewin Sime's biography of Meitner. Describing the 1941 meeting between Heisenberg and Niels Bohr, made famous by Michael Frayn's play *Copenhagen,* Aczel highlights documents released by the Bohr archive in 2002 suggesting that German scientists may have been more enthusiastic about the bomb than was once thought. He also points to evidence made public in 2005 that the United States dropped the bomb on Japan while key Japanese figures appeared interested in suing for peace, and that the bombings may have been part of a wider strategy to deal with an anticipated threat from the Soviet Union.

Aczel's phrasing can be hyperbolic: he refers to the first time that Fermi's CP-1 reactor was taken to criticality as "the most dangerous experiment in history". In *Atomic Awakening*, James Mahaffey makes clear that there have been many risky nuclear experiments since.

Mahaffey, a nuclear engineer and consultant, offers a long overview of nuclear research from the 1890s to the cold war. He writes for people who know the basics — testing the reader, for example, by introducing the Polish-born Marie Curie under her original name of Manya



The aborted Nuclear Engine for Rocket Vehicle Application mission to Mars was a novel attempt to harness nuclear energy.

Skłodowska. His description of research before the Second World War is informative. But it is his discussion of US cold-war nuclear research that is most special.

Much of the research undertaken in this age of wild experimentation was risky and harmful, Mahaffey admits, such as the irradiation experiments on plants and animals conducted at the Georgia Nuclear Aircraft Laboratory in the 1950s. But he also wants his readers to see the fun in going into the desert and firing up nuclear rockets and aircraft engines in total secrecy. He describes, for example, the bold tests of nuclear-powered ramjets pursued at Jackass Flats, Nevada, in the 1960s, followed by the remarkable series of nuclear rocket motors for spaceflight, including the Phoebus 2A, which Mahaffey calls the most powerful nuclear reactor ever built.

Mahaffey's US perspective sometimes leads him astray. He reports that "It has been over 60 years since anyone died in a criticality accident"; in fact, two workers died in Japan in the Tokaimura accident of 30 September 1999.

**Uranium Wars: The Scientific Rivalry that Created the Nuclear Age**
by Amir Aczel
Palgrave Macmillan: 2009.
256 pp. £18.99

**Atomic Awakening: A New Look at the History and Future of Nuclear Power**
by James Mahaffey
Pegasus Books: 2009. 368 pp. $26.95

And his claim that work on breeder reactors stopped in 1972 with the closure of the Fermi-1 reactor in Michigan is rectified when he discusses continued work in France.

*Atomic Awakening* taught me a great deal. It is persuasive and based on deep research, even if it contains few references. Mahaffey packs his book instead with personal observations, recollections and occasionally quirky footnotes. Yet, despite its subtitle, he has little to say about the future of nuclear power.

Both books contain errors. Aczel misnames plutonium "Pt-239" rather than Pu-239, and speaks of "refinement" rather than the more technical term 'enrichment'. And Mahaffey confuses the uranium oxide known in the industry as yellowcake with uranium tetra-fluoride, known as green salt.

Readers with a background in physical sciences or engineering would gain much from *Atomic Awakening*. For a non-scientist, *Uranium Wars* tells a good story well. ∎

**William J. Nuttall** is a senior lecturer in technology policy at Judge Business School, University of Cambridge, Cambridge CB2 1AG, UK, and author of *Nuclear Renaissance — Technologies and Policies for the Future of Nuclear Power*.
e-mail: wjn21@cam.ac.uk

# Valuing the digital economy

**Wired for Innovation: How Information Technology is Reshaping the Economy**
Erik Brynjolfsson and Adam Saunders
MIT Press: 2009. 134 pp. $18.95, £14.95

Information technology has clearly had an impact on our daily lives. It has also altered the global economy. But how can you measure the worth of a web search or an electronic file format? In this slender volume, Erik Brynjolfsson and Adam Saunders examine the links between innovation, productivity and value.

*Wired for Innovation* contrasts information with tangible goods. Unlike a physical product, information is not limited by its transportability or the number of times it can be used. Yet a consumer who uses a web search to book a holiday, for example, adds real value to the economy — and more than 8 billion web searches are made each day in the United States alone. Forming a demonstrable relationship between web activities and dollars generated is becoming increasingly important.

To illustrate why traditional economic measures are inadequate, Brynjolfsson and Saunders examine the impact of web-based products that seem, to the consumer, to be free, such as Google's search engine or Adobe's portable document format. These, they explain, follow two-sided markets: many search engines receive funding from keyword-based advertising, and Adobe makes money from selling other software products that use its file format. But the connections between the two sides of the business are invisible. Because Adobe doesn't sell its Reader, the contribution of that software to a traditional metric such as gross domestic product cannot be determined.

However, networked technologies also offer new ways to manage and analyse economic data at an increasingly fine-grained scale. The authors propose a tool kit for the emerging economics of the digital business world. Different devices are needed to monitor and test the market as it evolves. Some of their suggested indicators are not new, such as customer surplus — the aggregate net benefit that consumers receive from an item or service above what they paid for it — but they take on a new role in the digital environment.

The combination of mobile phones, global positioning systems, online maps and restaurant reviews, for example, has opened up



Web-based wonders: Google founders Sergey Brin (left) and Larry Page (right) and chief executive Eric Schmidt.

possibilities for revenues whose value exceeds that for each technology alone. Measuring that surplus requires further technologies to track links in the economic decision chain, such as smart 'sociometric' badges that monitor interpersonal interactions. Such developments would allow us to monitor flows of information through the economic and social system at a detailed level.

The human side of innovation must also be understood. Brynjolfsson and Saunders explain that information systems are meaningless without the ability to manage and accrue benefits from them. They propose "pillars" that should underpin a successful digital organization — where people are supported by technology, rather than the other way around. Open information access, rather than an excessively protective communication policy, is cited as one way for an organization to draw benefit from the human skills available to it — what you can't see, you can't extract value from.

Future economists and computer scientists may look to this decade as the beginning of a digital global economy. Small, light and portable, like the technologies it describes, *Wired for Innovation* will form a valuable reference for scholars searching for markers of this fundamental change. Those wishing to become a formative part of the new economy — rather than merely a footnote — should read this book. ∎
**John Gilbey** lectures in computer science at Aberystwyth University, UK.
e-mail: gilbey@bcs.org.uk

# Tips from the top of the career ladder

**Beyond the Boys' Club: Strategies for Achieving Career Success as a Woman Working in a Male-dominated Field**
Suzanne Doyle-Morris
Wit and Wisdom Press: 2009.
298 pp. £13.99, $22.99

As an executive coach who works with companies to recruit, retain and develop their female talent, Suzanne Doyle-Morris has long had an interest in issues concerning women in male-dominated professions, including academia and science. In her book, she offers advice to anyone who wishes to "understand the boys' club and move beyond it".

Doyle-Morris interviewed 21 senior women from a range of backgrounds to discover what made them survive and thrive in their workplaces. Her interviewees, spanning a wide age group, included a senior diplomat, a retired archaeologist who led digs in Iraq, and academics such as Athene Donald, professor of experimental physics at the University of Cambridge and 2009 recipient of a L'Oréal/UNESCO Women in Science award.

The difficulties women face in rising through their professions, Doyle-Morris argues, stem largely from old career structures. Too often it is assumed that an employee, usually male, has a partner, usually female, who will attend to domestic responsibilities, leaving the employee free to focus on work. As more women enter the workforce, they realize that this model does not suit them and they are made to feel inadequate if they do not play by the rules of this game.

Moreover, career advancement is not based entirely on merit. "People who get ahead are those who make others aware of their wins, and those who spend time developing relationships and their own personal profile," Doyle-Morris explains. As 80% of consumer decisions are said to be made by women, it makes good business sense for companies to put more women in senior positions. But even so, the rise to the top can be slow, especially in the science and engineering professions. In the United Kingdom, Doyle-Morris says, about 70% of women holding degrees in science, engineering or technology

M. PELLETIER/L'ORÉAL

do not currently work in these fields.

In addition to focusing on how individual women can develop strategies for success, *Beyond the Boys' Club* highlights effective government legislation. In Norway in 2002, for example, the government proposed legislation requiring 40% female representation on companies' executive boards, and by the deadline of 2008 the target had been reached.

Quotas and affirmative-action policies are often frowned upon by sceptics, who argue that this will lead to a loss of quality. But Doyle-Morris points out that men have always benefited from favours within male networks. "Most
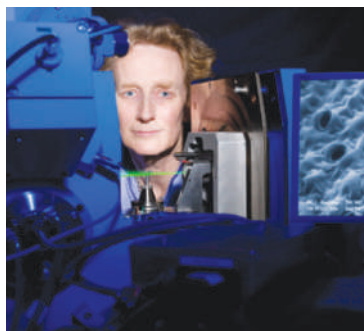
**Career success: physicist Athene Donald.**

men know that building relationships for mutual benefit is the only way to build a career," she notes, so women should not be afraid to do the same.

The book explains how best to develop these relationships. It offers tips on how to raise your profile, build your image, network within and outside the organization, take appropriate risks, negotiate office politics and choose a mentor or coach.

Although Doyle-Morris addresses most of these strategies to women working in companies, two professors add advice from academia. They discuss mentoring in universities and the use of flexible working time in

their laboratories. They suggest specializing in a field other than your PhD topic during the postdoc years, and finding cross-disciplinary collaborations. However, they don't address more particular problems, such as how to complete your doctorate if you suffer sexual harassment from a supervisor, or how to maximize visibility for your work in order to gain an influential position such as dean or president of a university.

Doyle-Morris has great experience and a passion for her subject. I hope she will write a sequel offering strategies for women in academic science — and that the sequel will offer insights from women belonging to different class backgrounds, races, nationalities, sexual orientations and physical abilities. ■

**Asha Gopinathan** is a researcher in neuroscience based in Trivandrum, India.
e-mail: dendron.15@gmail.com

# Q&A: The algorist

J.-P. HÉBERT

Having moved from engineering to art, **Jean-Pierre Hébert** applies mathematical rules to generate artworks that explore themes of chaos and determinism. As resident artist at the Kavli Institute for Theoretical Physics, and with an exhibition on in Los Angeles, Hébert explains his interest in algorithms.
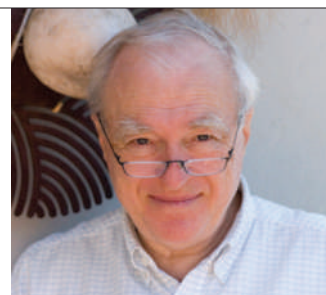
**How did you become fascinated by algorithms?**
At 19, I was exposed to the first computer that came to Europe. I took a summer job at IBM and became interested in mathematics and computer science. My taste in art was influenced by Mondrian and the Bauhaus. I was very interested in geometry. Soon I made algorithmic pieces by hand. Then I figured that I could program that.

J.-P. HÉBERT

**Some of your works are computer-drawn, others are sculptures. Which do you prefer?**
I have more affinity with the physical than the virtual world. I like to mark paper or water or sand, and I like to explore algorithms without using computers. There is a richness in something beautifully made on paper, stone, clay or canvas that is more inspiring than a flat image.

**How have you interacted with researchers?**
I've always had a strong interest in physics. In collaborations with scientists I learnt about new phenomena in physics and nature and about new ways to express them. I have learnt tricks to define spirals and made graphical presentations of the gravitation around black holes.

**How did your style develop?**
At first my work was monochromatic and flat, and used only certain themes in geometry. I wanted to animate it more and make it happen in three dimensions — to make it subject to fields and forces and waves, and to submit it to phenomena like reflections on the surface of water.

**How did you define the 'algorist' movement?**
We established, 40 years ago, that the work should

be based on algorithms but it should also produce an object of art — something concrete, such as a drawing, a painting, a sculpture, a piece of music. It should not be simply the concept behind an algorithm.

**Are algorithms important in art?**
Over the past century most artists were algorists, even if they did not know it. Mondrian has an algorithm; cubism was a set of algorithms. Algorithms are just a tool, as is a computer, a brush or a pencil. The personality of the artist takes over and should transpire through the work. Algorithms are not an end in themselves. ■
Interview by *Nature* reporter **Daniel Cressey**.

**Jean-Pierre Hébert: Drawings as Thoughts**
SCI-Arc Library, Southern California Institute of Architecture, Los Angeles
until 13 December 2009

Kavli Institute for Theoretical Physics, University of California, Santa Barbara
A sand sculpture has been placed on permanent display.

# NEWS & VIEWS
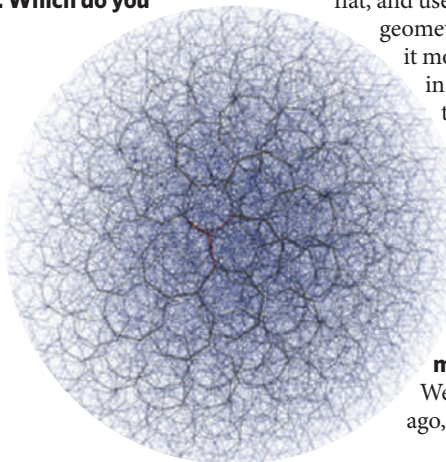
**DRUG DISCOVERY**

# Predicting promiscuity

Andrew L. Hopkins

**Computational methods that reliably predict the biological activities of compounds have long been sought. The validation of one such method suggests that *in silico* predictions for drug discovery have come of age.**

The metaphor of drugs acting as 'magic bullets', selectively binding to specific physiological targets, dates back to Paul Ehrlich and the foundation of modern drug discovery. But in recent years, the observation that drugs often bind to more than one molecular target — that they exhibit polypharmacology — has gained attention[1], suggesting that a 'magic shotgun' analogy might be more apt (Fig. 1). To fully understand the actions of a drug, knowledge of its polypharmacology is clearly essential. On page 175 of this issue, Keiser *et al.*[2] report a computational tool that generates predictions of the pharmacological profile of drugs, and provide experimental validation of their method.

The interactions of drugs with off-target proteins have conventionally been viewed as undesirable 'promiscuity', responsible for unwanted side effects. But in many cases — ranging from certain older psychiatric drugs to modern anticancer therapies — this promiscuity is intrinsic to the drug's therapeutic efficacy. An understanding of polypharmacology can therefore explain why some drugs work better than expected, or why different drugs that supposedly act on the same target vary in their side effects. Perhaps more importantly, such insight offers opportunities to develop drugs for diseases that don't yet have therapies. But screening drugs against all the proteins expressed by the 21,000-plus genes in the human genome

is currently unfeasible. A long-time goal of chemoinformatics has therefore been to develop computational techniques that predict the proteins to which drugs are likely to bind.

Various *in silico* methods for predicting the pharmacological profile of drugs are in development, the most well known of which is to 'dock' the three-dimensional structure of a compound virtually into the structure of a protein. But among the limitations of docking methods is the need for high-resolution X-ray crystal structures of proteins. These are particularly difficult to obtain for membrane-bound proteins, which account for 60% of drug targets. An alternative approach has therefore been developed that does not require protein structures. This approach works by analysing the chemical structures of ligand molecules that are known to bind to drug targets, to identify the structural motifs responsible for the binding.

Several groups are also developing machine-learning methods that identify molecular binding motifs that might relate compounds that act at one protein to those that act at another, as a means of inferring the unknown polypharmacology of a drug. These methods are often based on Bayesian statistical analyses[3]. Keiser and colleagues' technique[2] is also ligand-based, but it uses a new algorithm to build up profiles of the patterns (or fingerprints) of chemical structures for all the reported ligands of a

drug target. These profiles can then be used to calculate the probability that another structurally unrelated compound contains the same chemical patterns[4]. The underlying mathematics of this similarity ensemble approach[2] (SEA) is inspired by the methods used to compare the similarity of amino-acid sequences in proteins.

Keiser *et al.* used their technique to analyse more than 3,600 compounds, each of which was either an approved drug or a compound that had been investigated as a possible drug. They computationally screened the chemical structures of these compounds against databases of hundreds of thousands of biologically active compounds, for which the binding profiles to a panel of more than 1,400 proteins are known. This yielded a massive network of predicted drug–target interactions for the 3,600 compounds in question. Nearly 7,000 of these predicted interactions had sufficiently high probabilities to warrant further study.

The authors investigated the most likely predictions of drug polypharmacology for several compounds, both retrospectively (by database and literature searching) and prospectively (by *in vitro* and *in vivo* screening). Although only a tiny fraction of all the predictions was tested, the authors' initial validation data confirmed that two out of every three high-confidence predictions were correct. This implies that the polypharmacology network between drugs
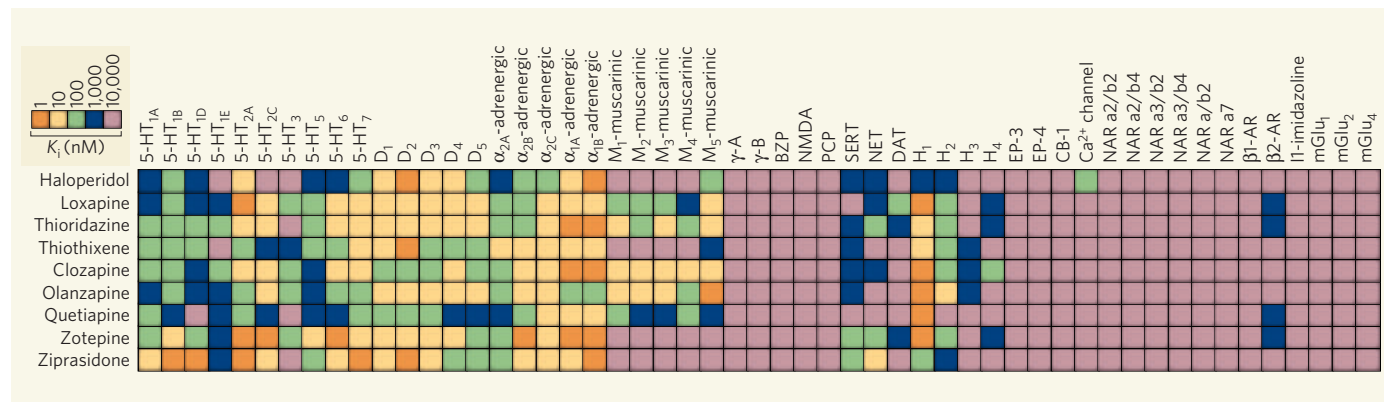


**Figure 1 | One drug, many targets.** Drugs often exhibit polypharmacology — they bind at more than one molecular site. The chart shows the known affinity ($K_i$) values of several antipsychotic drugs for a panel of receptors. Low $K_i$ values correspond to high affinities. Keiser *et al.*[2] report and validate a computational method that predicts the molecular sites to which a drug will bind. (Adapted from ref. 1.)

and drug targets is far denser than is currently suggested by the published data[5] (see Fig. 1 on page 176).

Several of Keiser and colleagues' predictions[2] provide fascinating insights into the previously unknown pharmacology of well-known drugs. For example, the antidepressants fluoxetine (Prozac) and paroxetine (Paxil) are selective serotonin reuptake inhibitors (SSRIs) — they work by inhibiting the uptake of the neuro-transmitter serotonin into presynaptic neurons. But the authors' computational method strongly predicted that these drugs also act as beta-blockers (they bind to β-adrenergic receptors in, for example, blood vessels and heart muscle), a prediction that was confirmed *in vitro*. This result suggests that some of the side effects that occur when patients stop taking SSRI antidepressants might be caused by a rebound in β-adrenergic signalling.

As serotonin and β-adrenergic receptors bind to chemically similar, naturally occurring ligands, the beta-blocker activity of SSRIs might not be too unexpected. A much tougher challenge is to predict a compound's polypharmacology for drug targets that share no discernible similarity in terms of amino-acid sequence, protein structure or endogenous ligands. Again, Keiser and colleagues' technique comes up with the goods: the authors confirmed the SEA prediction that delavirdine (Rescriptor, an HIV-1 reverse transcriptase inhibitor) also binds to histamine $H_4$ receptors. These drug targets have no biological similarity whatsoever, yet the structure of delavirdine 'encodes' its ability to recognize both reverse transcriptase and $H_4$ receptors. The binding of delavirdine to these receptors might explain why it sometimes causes painful skin rashes as a side effect.

Although computational techniques for drug discovery have been around for some time, confidence in their use has often been lacking. But with the growing availability of free databases of pharmacological information, such as ChEMBL[6] and PubChem[7], the field of chemoinformatics finally has sufficiently large quantities of data for machine-learning methods to be developed that could reliably predict drug activity. Keiser and colleagues' experimental validation[2] of their SEA for predicting polypharmacology from chemical structure shows that such methods are truly beginning to bear fruit. ∎

Andrew L. Hopkins is in the Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK.
e-mail: a.hopkins@dundee.ac.uk

1. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. *Nature Rev. Drug Discov.* **3,** 353–359 (2004).
2. Keiser, M. J. *et al. Nature* **462,** 175–181 (2009).
3. Nigsch, F., Bender, A., Jenkins, J. L. & Mitchell, J. B. *J. Chem. Inf. Model.* **48,** 2313–2325 (2008).
4. Keiser, M. J. *et al. Nature Biotechnol.* **25,** 197–206 (2007).
5. Yildirim, M. A. *et al. Nature Biotechnol.* **25,** 1119–1126 (2007).
6. www.ebi.ac.uk/chembl
7. http://pubchem.ncbi.nlm.nih.gov

# A fossil record for exoplanets

Marc Pinsonneault

**Stars that host planets experience more mixing of their internal elements than do stars that lack such companions. This correlation may serve as a useful diagnostic in the search for planets around stars other than the Sun.**

The planets that have been discovered beyond our Solar System — exoplanets — are extremely diverse. As the field of exoplanetary research has matured, systems containing multiple planets have been detected and classified, and our detection sensitivity has extended to planets with lower masses and longer orbital periods. Planet formation is highly sensitive to the properties of the accretion disks of gas around the protostars, and the masses and lifetimes of these disks are known to vary dramatically from system to system. The discovery that very different types of exoplanetary system exist is therefore not surprising, and competing theoretical models for their formation have emerged. However, evidence of the formation process around stars such as the Sun is typically erased by the passage of time, and directly correlating planetary properties with formation scenarios has not been possible. In this issue, Israelian *et al.*[1] (page 189) demonstrate that the abundance of lithium — a diagnostic of the internal mixing of elements in stars — differs significantly between stars that host planets and those that do not.

The depletion of lithium on the surface of a star is a sign either that the matter at the surface layers has been exposed to temperatures greater than 2.5 million kelvin or that the lithium has settled — like sugar in an unstirred glass of iced tea — below the star's surface. The latter process has a long timescale for Sun-like stars. Standard stellar-evolution theory predicts that protostars have a thick, turbulent, convective envelope of gas with a temperature at its base that is sufficiently high to burn some, but not all, of the lithium. This envelope retreats as the star stabilizes and reaches a Sun-like 'main-sequence' state, in which energy is created through the fusion of hydrogen at the star's core. Lithium depletion is progressively more severe at all stages for lower-mass stars and those with a higher concentration of heavy elements. However, no main-sequence depletion of lithium is predicted for stars more massive than about 0.9 solar masses, and Sun-like stars are observed to arrive at the main-sequence state with relatively high surface lithium abundances.

This basic picture of the thermal structure of stars is confirmed by seismic measurements of the thickness of the Sun's convection zone, which is too thin and cold at its base to destroy lithium[2]. However, the Sun is depleted in lithium by a factor of more than 100 relative
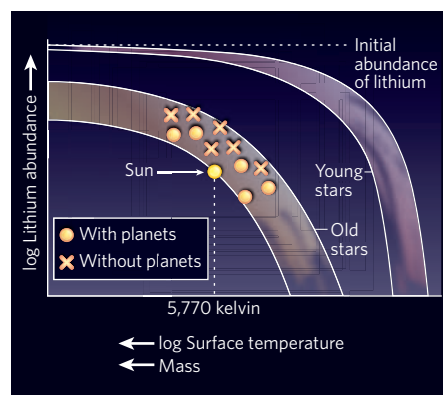


**Figure 1 | Stellar lithium depletion.** Lithium is depleted in stars early in their evolution, when they have a thick, turbulent, convective envelope of gas; turbulence cycles the fragile element to the stars' centre, where it is destroyed by nuclear reactions. This turbulent mixed envelope becomes thinner once hydrogen fusion is established in the stars' cores (at which point they become 'main-sequence' stars), but young stars retain a record of their early history (top band in this schematic illustration). The lithium depletion rate in this early stage increases with decreasing mass and surface temperature. After the main-sequence state is reached, stellar rotation can further enhance lithium depletion by slowly mixing it into the stars' interiors. This extra degree of mixing varies from star to star as they age, yielding a dispersion in lithium abundances in old stars (bottom band). Israelian *et al.*[1] find that the Sun and Sun-like stars that host planets (circles) have lower lithium abundances than those that do not (crosses). This indicates that planet formation must have influenced the lithium depletion process.

to the value measured in meteorites, and by a factor of more than 50 relative to young Sun-like stars[3]. This is a general phenomenon: we observe that, for Sun-like stars, lithium abundance declines with age during the main-sequence stage and, furthermore, that the degree of depletion differs for stars of the same mass, composition and age[4]. An additional physical mechanism is thus required to explain this complex observational pattern.

Israelian and colleagues[1] demonstrate that lithium depletion is more severe in Sun-like stars that host planets than in those that do not (Fig. 1). They choose a sample of stars with similar surface temperatures, and therefore masses, so as to minimize the known mass dependence of lithium depletion. They find no convincing evidence that the relative

depletion trends can be explained by differences between the two types of star in either age or the abundances of heavy elements such as iron. This discovery provides both a valuable method for screening possible planet-hosting systems and a potential window on the planet-formation process.

The authors[1] do not provide direct evidence for the underlying cause of the differences they have discovered. Stellar rotation is an attractive explanation, however. Rotation can induce both turbulence and global circulation currents that can mix lithium into the interior regions, where it can be destroyed. Stars are born with a range of rotation rates, leading naturally to a range of depletion histories. But testing this theory directly has proved difficult. Low-mass stars such as the Sun have magnetized winds that cause the stars to spin down as they age, and these winds cause rapid rotators to spin down preferentially relative to slow ones[5,6]. At late ages, evidence of the initial rotation is erased, leaving only potential fossil evidence, such as the record of lithium depletion. Highly precise rotation-period measurements in coeval stars, such as those in star clusters, may eventually permit this hypothesis to be tested directly.

The total angular-momentum content of a star, which describes its rotational state, is determined by interactions between the protostar and its accretion disk. The formation of planets in such disks could therefore plausibly modify the rotation of the host star. Protostars emerge from their cocoons of gas rotating at a modest rate, despite the fact that the parent gas clouds must have collapsed and spun up by enormous factors during the star-formation process. The most commonly accepted reason for the star's relatively modest initial rotation involves the transport of angular momentum, via a magnetic field, within a circumstellar accretion disk, and between the disk and the parent star[7]. Slow rotation can be explained if angular momentum is extracted from the incoming gas before it is incorporated into the star. Once a star is mostly assembled, it contracts and would ordinarily spin up. However, in the presence of a coupled accretion disk, the star can be made to co-rotate with the inner edge of the disk, preventing it from spinning up until it can become decoupled from the disk.

The observed wide range in stellar rotation rates would be explained by variations in the initial gas-accretion rate and the lifetime of the disks. The formation of giant planets can induce gaps in the accretion disk, disrupting the magnetic star–disk connection and thus plausibly permitting the star to spin up. The next stage in understanding Israelian and colleagues' results[1] is to examine rotation and to look for planets in young stars, which retain a memory of their natal rotation. Although technically challenging because of the associated high variability of young stars, such studies will be necessary to establish a physical explanation for the intriguing link between the lithium abundances described by Israelian et al. and the planet-formation process. ■

Marc Pinsonneault is in the Astronomy Department, Ohio State University, Columbus, Ohio 43210, USA.
e-mail: pinsono@astronomy.ohio-state.edu

1. Israelian, G. et al. Nature 462, 189–191 (2009).
2. Basu, S. & Antia, H. M. Mon. Not. R. Astron. Soc. 287, 189–198 (1997).
3. Lodders, K. Astrophys. J. 591, 1220–1247 (2003).
4. Sestito, P. & Randich, S. Astron. Astrophys. 442, 615–627 (2005).
5. Skumanich, A. Astrophys. J. 171, 565–567 (1972).
6. Kawaler, S. D. Astrophys. J. 333, 236–247 (1988).
7. Shu, F. et al. Astrophys. J. 429, 781–796 (1994).

## LANGUAGE EVOLUTION

# The importance of being human

Martin H. Dominguez and Pasko Rakic

**The *FOXP2* gene is implicated in the development of human speech and language. A comparison of the human and chimpanzee FOXP2 proteins highlights the differences in function in the two species.**

Some genes find instant favour in scientific culture and, like popular celebrities, remain under the close watch of their devoted following. *FOXP2* has many of the qualities of an in-vogue gene — relevance to human disease, evolutionary significance and a prominent role in brain development. The gene was first found to be mutated in a family with language dysfunction[1], hence its implication in the development and evolution of human speech and language, and its reputation as the 'language gene'. Examination of the human *FOXP2* gene and protein sequences[2] reveals accelerated evolution of its amino-acid sequence (specifically, two new amino-acid substitutions) since humans' common ancestor with chimpanzees, suggesting that the gene has a prominent role in the acquisition of language in humans. But like the leading character in Oscar Wilde's play *The Importance of Being Ernest*, *FOXP2* leads a double life — on the one hand, it has a starring role as the language gene, but on the other hand it has roles that are less glamorous, although still worthy of scientific attention. On page 213 of this issue, Konopka et al.[3] advance our knowledge of *FOXP2* by publicizing some

of its private doings — the underpinnings of its gene-regulatory networks, especially those conferring functions that are unique to humans.

*FOXP2* is present in vocal and non-vocal animals, is important for muscle coordination in vocalization, but is also associated with many other motor functions in all vertebrate species examined. The most obvious consequence of loss of function of FOXP2 in humans[1] and rodents[4] is impairment of motor skills and coordination. Problems in motor sequencing actions or procedural learning (the acquisition of fine motor skills), including those related to the mouth and face, thus can manifest as disorders of speech and language[5]. The involvement of *FOXP2* in learning corroborates nicely with quantitative changes in synaptic plasticity (changes in the strength of a synapse between neurons) that are seen in the relevant mouse neural circuits when the gene is mutated[4] or replaced with the human version[6].

Impairment of motor-learning skills can have profound effects in humans, in whom a loss-of-function *FOXP2* mutation[1] causes disturbances in language comprehension,

grammar and syntax. Such uniquely human aspects of language processing and delivery may be the result of selective elaboration of particular anatomical regions, or the acquisition of novel functions by those regions during human evolution[7]. Indeed, *FOXP2* exhibits region-specific expression in the developing human cerebral cortex, where its expression is highest in cortical regions, such as the orbital prefrontal and perisylvian association areas, thought to be responsible for higher cognitive functions and language[8]. *FOXP2* is also strongly expressed in the human striatum[9], a site in the brain that is involved in cognition and motor coordination, and changes in the branching pattern of dendrites on neurons at this site are observed[6] in mice genetically engineered to express the human version of *FOXP2*.

Complex traits such as language undoubtedly require an equally complex nexus of evolved pathways. To further understand what FOXP2 does on a molecular level, two articles[8,10] have revealed some of its probable targets, but neither study compared the regulatory effects of human and ancestral FOXP2. This is precisely what Konopka and colleagues[3] have done, using whole-genome arrays to detect differences in gene expression in human neuronal cell lines expressing either human FOXP2 (FOXP2^human) or the ancestral protein, FOXP2^chimp. The authors find that a substantial number of FOXP2 target genes are differentially regulated by FOXP2^human and FOXP2^chimp. Many of these genes met the criteria for positive selection during human evolution (although the authors had no way of assessing their statistical significance).

This places their findings in harmony with previous results[9–11] that show FOXP2-related genes as evolutionary arbiters. Because the authors examine human-specific gene regulation by FOXP2, their work[3] may provide our first window on the co-evolution of regulatory networks that are important for human-specific features such as language, which probably require a number of genetic changes working in concert. So far, a variant of one FOXP2 target gene, *CNTNAP2*, has been associated[11] with heritable language impairment in children. Additional work is essential to elucidate the relationships between *FOXP2* and the many other genes involved in these pathways.

Konopka *et al.*[3] performed mass spectrometry comparisons of FOXP2[chimp] and FOXP2[human] immunoprecipitates, but did not find major differences in the composition of associated proteins, which might have been expected if they interact with different transactivating factors. Furthermore, the region of the FOXP2 protein at which the two human-specific amino-acid changes occurred has an unknown function, leaving open the possibility that different DNA-binding characteristics might account for the observed disparities in gene expression. An alternative explanation for differential target-gene regulation by FOXP2[human] and FOXP2[chimp] may be the higher level of the human protein in the tested cells compared with the ancestral protein, perhaps owing to greater protein stability — a larger pool of free FOXP2 might permit a relaxed DNA-binding specificity. However, because FOXP2[chimp] causes greater changes in gene expression than FOXP2[human], such an alternative mechanism is less likely. Instead, the DNA that surrounds and includes the FOXP2-binding sites in the regulatory regions of target genes may provide a genomic context for their recognition by FOXP2[human].

To explore this possibility, the authors[3] chose eight differentially expressed genes, and found that, for six of them, their short promoter regions exhibited differential activity that paralleled the quantitative expression data. A careful analysis of those promoter DNA sequences may reveal common elements that result in disparate activities with FOXP2[human] and FOXP2[chimp]. Co-evolution may likewise be important in the recognition sites, as similar sequences across the human genome may bear evidence of human-lineage-specific alteration.

In summary, Konopka and colleagues' work[3] does what important discoveries usually do: it answers many questions, but raises even more. It confirms evolutionary relationships between FOXP2 proteins in different species revealed through genome sequencing, and uncovers potential mechanisms underlying the elaboration of human-specific traits such as speech. However, it also provides a starting point for future studies of the molecular basis of language and human evolution. ■

Martin H. Dominguez and Pasko Rakic are in the Department of Neurobiology and the Kavli Institute for Neuroscience at Yale University, New Haven, Connecticut 06510, USA.
e-mail: pasko.rakic@yale.edu

1. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. *Nature* **413,** 519–523 (2001).
2. Enard, W. *et al. Nature* **418,** 869–872 (2002).
3. Konopka, G. *et al. Nature* **462,** 213–217 (2009).
4. Groszer, M. *et al. Curr. Biol.* **18,** 354–362 (2008).
5. Watkins, K. E., Dronkers, N. F. & Vargha-Khadem, F. *Brain* **125,** 452–464 (2002).
6. Enard, W. *et al. Cell* **137,** 961–971 (2009).
7. Rakic, P. *Nature Rev. Neurosci.* **10,** 724–735 (2009).
8. Johnson, M. B. *et al. Neuron* **62,** 494–509 (2009).
9. Spiteri, E. *et al. Am. J. Hum. Genet.* **81,** 1144–1157 (2007).
10. Vernes, S. C. *et al. Am. J. Hum. Genet.* **81,** 1232–1250 (2007).
11. Vernes, S. C. *et al. N. Engl. J. Med.* **359,** 2337–2345 (2008).

**CONDENSED-MATTER PHYSICS**

# Dirac electrons broken to pieces

Alberto F. Morpurgo

**Graphene continues to surprise physicists with its remarkable electronic properties. Experiments now show that electrons in the material can team up to behave as if they are only fragments of themselves.**

The fractional quantum Hall effect (FQHE) is a fascinating form of collective electronic behaviour. It arises when electrons in a strong magnetic field — applied at a right angle to the plane in which the electrons flow — act together to behave like particles with a charge that is a fraction of an electron's charge[1]. Its observation requires the use of two-dimensional systems virtually free of disorder. This is why, since its discovery by Daniel Tsui and Horst Störmer in 1982, the effect has been studied in ultrapure semiconductor hetero-structures (devices that contain thin layers of one or more semiconductors) grown in an ultrahigh vacuum. In this issue, Du *et al.*[2] (page 192) and Bolotin *et al.*[3] (page 196) show that the FQHE can also be observed in graphene — a one-atom-thick sheet of graphitic carbon, the production of which requires no more sophistication than a common adhesive tape to manually exfoliate graphite in ambient conditions[4].

In the presence of a magnetic field (*B*), electrons carrying current in a bulk material are subjected to a force (Lorentz force) that deflects them sideways in a direction perpendicular to both the applied field and the current. The deflected electrons accumulate at the edges of the material and generate an electric field that exactly compensates for the Lorentz force. The resulting voltage causes a finite electrical resistance — termed the Hall resistance ($R_H$) — that, according to classical physics, increases linearly with the strength of the magnetic field, but is otherwise featureless. This unassuming electronic behaviour is what Edwin Hall first reported[5] in 1879.

It was 100 years before Klaus von Klitzing discovered[6] that, in two-dimensional conductors, the Hall effect comes with a twist. The dependence of $R_H$ on *B* is far from featureless: it exhibits a series of pronounced plateaux at values precisely equal to $R_H = h/ne^2$, where *e* is the electron charge, *h* is Planck's constant and *n* is an integer number — which explains why the phenomenon is dubbed the integer quantum Hall effect. The plateaux originate from the quantization of the motion of individual electrons in a magnetic field, which gives rise to a ladder of energy levels known as Landau levels.

In the centre of the conductor, the Landau levels are separated by 'forbidden' energy gaps in which no electronic quantum state is available. At the edges, however, the Landau levels bend upwards and the gaps close, with each of the Landau levels forming a channel in which electrons can propagate in a single direction. When the Fermi level — the highest energy level occupied by the electrons — is located between two Landau levels in the centre of the system, electronic transport occurs through these 'edge states', and the quantum Hall effect occurs. The integer *n* indexing the plateaux of the Hall resistance corresponds to the number of edge states occupied by electrons[7], and the quantization of the resistance reflects the fact that the smallest unit of charge that can propagate in such states is exactly the electron charge *e*.

In the FQHE, *n* becomes a fraction[1]. This seemingly minor difference has a deep conceptual significance, as first realized by Robert Laughlin[8]. It indicates that the smallest unit of charge in the system is not the electron charge, but a fraction of it. The effect is due to correlations induced by strong interactions between the electrons: it looks as if electrons in the fractional quantum Hall regime spontaneously break into pieces or 'quasiparticles' — a counter-intuitive idea that has proved correct experimentally. But this is only one of the weird aspects of the FQHE, because these fractionally charged quasiparticles are expected to have other unusual properties, such as their odd quantum statistics — that is, the way in which the quantum-mechanical wavefunction of the system changes when two quasiparticles are swapped.

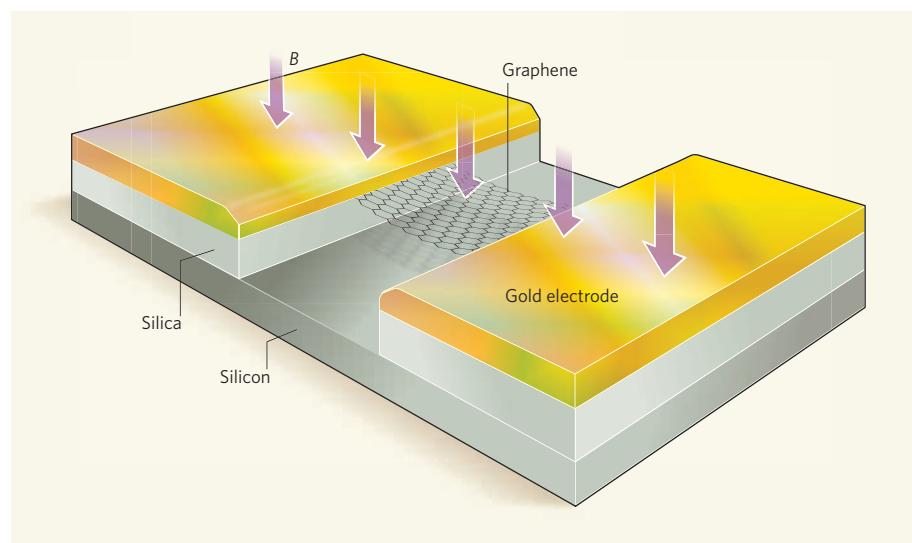At the end of 2005, the groups of Andre Geim

**Figure 1 | Suspended graphene.** Du *et al.*[2] and Bolotin *et al.*[3] observed a form of collective electronic phenomenon, known as the fractional quantum Hall effect, in graphene — a one-atom-thick honeycomb-like lattice of carbon — by exposing it to a strong magnetic field (*B*). The experimental set-up, which was pivotal for the observation of the effect, involved suspending a single sheet of graphene, supported by two gold electrodes, above a composite silicon/silica substrate.

and of Philip Kim investigated the magneto-resistance (a material's ability to change its electrical resistance when placed in a magnetic field) of graphene[9,10] that was obtained by peeling it from graphite using adhesive tape — a technique developed by Geim and colleagues[4] a year earlier. The authors[9,10] observed the integer quantum Hall effect and found that electrons in graphene effectively behave as massless relativistic particles, and that their dynamics are described by the (relativistic) Dirac equation. The properties of such Dirac electrons cannot be understood by a simple extension of the (non-relativistic) Schrödinger equation — which has been shown to describe the electronic behaviour of other condensed-matter systems — and require the introduction of new concepts. This startling discovery marked the start of graphene electronics, one of the most exciting research fields of the past decade.

The FQHE of Dirac electrons is an unexplored area, and its observation — now reported by Du and colleagues[2] and Bolotin and colleagues[3] — goes to show how fast the field is advancing. The experiments[2,3] rely on devices in which graphene is suspended (Fig. 1), allowing its isolation from the underlying substrate that would otherwise introduce substantial disorder and prevent the detection of the FQHE. The researchers pioneered the fabrication technique of suspended-graphene devices last year. With the demonstration of the FQHE for quasiparticles with a fraction $n = 1/3$ of the electron's charge, they are now seeing the efforts spent in developing the technique paying off. Because suspended-graphene devices are only a few micrometres in size, resistance measurements cannot be performed in the conventional device configuration used by Hall in his experiments[5], which relies on the use of

four electrical contacts. The researchers opted for a less conventional, two-contact geometry, which still proved sufficient to demonstrate the fractional quantization of the resistance. This strategy allowed the authors[2,3] to show that the effect in graphene is remarkably robust: the experiments revealed the $n = 1/3$ plateau up to a temperature of about 20 kelvin, a factor of 100 higher than that obtained in commonly used semiconductor heterostructures.

The results reported by Du *et al.*[2] and Bolotin *et al.*[3] thus convincingly demonstrate the occurrence of the FQHE, as well as the presence in graphene of the underlying electronic

correlations induced by electron–electron interactions. The remaining challenge is to detect new physical phenomena originating from the Dirac nature of electrons. Theories are being developed, and some already exist, that predict measurable effects. Graphene also opens new experimental routes for the study of the FQHE. In contrast to semiconductor heterostructures, in which the two-dimensional electron system is buried inside the device, electrons in graphene are exposed at the surface. This enables the use of scanning-probe-microscopy techniques — which rely on the interaction between a scanning probe and a specimen's surface to form an image — to investigate the nature of fractionally charged quasiparticles. Clearly, the time is ripe for more work on graphene electronics, both experimental and theoretical. Dirac electrons have just been broken into pieces. Putting the pieces of their physics together will take some time. ∎

Alberto F. Morpurgo is in the Department of Condensed Matter Physics and the Group of Applied Physics, University of Geneva, CH1211 Geneva, Switzerland.
e-mail: alberto.morpurgo@unige.ch

1. Tsui, D. C., Stormer, H. L. & Gossard, A. C. *Phys. Rev. Lett.* **48,** 1559–1562 (1982).
2. Du, X., Skachko, I., Duerr, F., Luican, A. & Andrei, E. Y. *Nature* **462,** 192–195 (2009).
3. Bolotin, K. I., Ghahari, F., Shulman, M. D., Stormer, H. L. & Kim, P. *Nature* **462,** 196–199 (2009).
4. Novoselov, K. S. *et al. Science* **306,** 666–669 (2004).
5. Hall, E. *Am. J. Math.* **2,** 287–292 (1879).
6. von Klitzing, K., Dorda, G. & Pepper, M. *Phys. Rev. Lett.* **45,** 494–497 (1980).
7. Büttiker, M. *Phys. Rev. B* **38,** 9375–9389 (1988).
8. Laughlin, R. B. *Phys. Rev. Lett.* **50,** 1395–1398 (1983).
9. Novoselov, K. S. *et al. Nature* **438,** 197–200 (2005).
10. Zhang, Y., Tan, Y.-W., Stormer, H. L. & Kim, P. *Nature* **438,** 201–204 (2005).

**CHEMICAL BIOLOGY**

# A Notch above other inhibitors

Paramjit S. Arora and Aseem Z. Ansari

**A tenet of drug discovery states that molecules greater than a certain size don't enter cells. But not only do certain synthetic peptides refute this idea, they also inhibit 'undruggable' biological targets.**

A basic limitation of drug development is the inability of traditional 'small-molecule' pharmaceuticals to target large protein interfaces, many of which are desirable drug targets. On page 182 of this issue, Moellering *et al.*[1] report that the solution could be to work with bigger molecules. They describe the design and synthesis of stabilized peptide helices that bind to a protein–protein interface in the Notch signalling pathway — a system that governs fundamental aspects of cell development, proliferation and death. Their remarkable results highlight the potential of molecules that

mimic the secondary structures of proteins to target normally intractable protein–protein interactions.

Most drugs are small molecules that target molecular pockets in enzymes or protein receptors. In general, such molecules fail to bind to the large, often flat, interfaces that hold protein partners together — or bind with insufficient specificity and affinity to be useful for pharmaceutical applications[2]. A broad effort is therefore under way in chemical biology to develop ligand molecules that target protein–protein interfaces. The focus for this is on screening

libraries of synthetic compounds in the hope of finding weakly binding ligands that might be good starting points for drug discovery.

A complementary strategy is to rationally design synthetic ligands by adapting nature's protein-recognition principles. This approach is based on the fact that the secondary structures of proteins have vital roles in their interactions with other biomolecules[3], so that stable mimics of these structures may provide potential ligands for protein interfaces. Moellering and colleagues' stabilized peptide helices[1] offer further proof-of-principle for this approach: they bind to the interface between the gene-regulatory protein (transcription factor) and the multi-protein co-activator that is required for transcription of genes in the Notch signalling pathway.

Inappropriate Notch function has been implicated in several human diseases, including cancers of the lung, ovary, pancreas and — of particular relevance to Moellering and colleagues' study — T-cell acute lymphoblastic leukaemia (T-ALL, a cancer of immature T cells). Unlike typical transcription factors, Notch is a transmembrane receptor that is displayed at cell surfaces, and binds to protein ligands on the surfaces of adjacent cells. Ligand binding to the extracellular domain of Notch triggers intramembrane cleavage of the receptor, mediated by the γ-secretase enzyme. The liberated intracellular domain of Notch (Notch[ICN]) then enters the cell nucleus, where it docks with the DNA-bound transcription factor CSL and helps to bind a co-activator, MAML1 (Fig. 1a). Recruitment of the co-activator and its associated transcriptional machinery initiates the expression of the Notch-targeted genes[4].

The binding domain of MAML1 adopts an α-helical conformation to target the protein–protein interface of CSL–Notch[ICN]. Moellering et al. constrained various peptide segments from this binding domain in stable helical conformations by crosslinking side chains of the amino acids that reside on the same face on the helix[5,6]. The authors found that a 'stapled helix' obtained from the wild-type sequence of the co-activator competitively inhibits MAML1 binding to CSL–Notch[ICN], and represses Notch-mediated gene expression dose dependently in a panel of T-ALL cell lines (Fig. 1b). More significantly, the stapled helix reduces leukaemic cell proliferation in a mouse model of T-ALL.

Strikingly, Moellering and colleagues' stapled helix bypasses some of the strict limitations that have been placed on drug discovery. When designing potential drug candidates, most medicinal chemists adhere to the Lipinski rules, which stipulate that the molecular mass of a drug should not exceed 500 daltons. However,
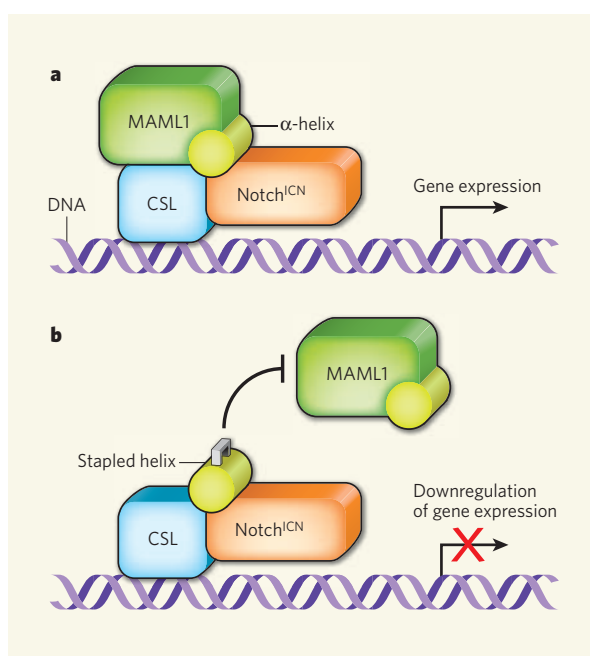


**Figure 1 | Inhibition of the Notch signalling pathway.** Notch proteins are transmembrane receptors found in cell membranes. When Notch's extracellular region binds to a ligand, the intracellular domain (Notch[ICN]) is released, and enters the cell nucleus. **a**, Notch[ICN] docks with CSL, a DNA-bound transcription factor. The CSL–Notch[ICN] complex is recognized by an α-helical structure in the binding site of MAML1, a co-activator protein. Binding of MAML1 to the CSL–Notch[ICN] complex initiates the expression of Notch-targeted genes. **b**, Moellering et al.[1] prepared peptide segments of the MAML1 binding site, constrained into α-helical conformations by hydrocarbon 'staples'. The stapled helices bind to the CSL–Notch[ICN] complex, preventing MAML1 from binding and so downregulating the expression of Notch-targeted genes.

the stapled peptides[1] are several times that size, and yet still traffic to the nucleus and efficiently compete with cellular transcription factors, contrary to the Lipinski rules' predictions. Several arginine amino-acid residues in Moellering and colleagues' helices might facilitate cellular uptake of the peptides, because polyarginines are cell permeable[7]. But there is accumulating evidence that the hydrocarbon crosslinks in stapled peptides might also favourably influence the cellular uptake of this class of compounds[8].

One of the main challenges for drugs that target transcriptional complexes is that they have to disrupt the cooperative assembly of several proteins on DNA. Cells fine-tune the balance of a transcription factor complex's concentration with its affinity for specific DNA sites, coordinating many weak individual interactions to stabilize the optimal assembly of proteins that regulates the expression of target genes. Targeting any single surface within a complex is therefore likely to have a limited effect on the function of the complex as a whole.

Another daunting hurdle is imposed by the ability of several physiologically unrelated transcription factors to interact with a common surface on a co-activator[4]. A minimally sized peptide that mimics such a surface is

therefore likely to perturb the function of many transcription factors. The fact that Moellering and colleagues' stapled helix[1] selectively blocks the Notch–CSL complex is particularly intriguing in this context. Its selectivity may derive from the ability of the stapled helix to dock into the Notch–CSL surface groove, thus inhibiting the binding of these transcription factors to MAML1, but not that of other transcription factors that bind to MAML1.

Moellering and colleagues' report[1] highlights the potential of synthetic ligands to target gene-specific transcription factors in general. The inappropriate activation of these key regulators of cell fate is linked to the aetiology of numerous genetic diseases, making them attractive drug targets. But although some of the most successful therapeutics act on a handful of transcription factors, including p53 and oestrogen receptors, the majority of transcription factors (more than 2,000 of them) have so far proved to be 'undruggable'[9,10].

Nevertheless, tantalizing hints are emerging about how to tackle these recalcitrant proteins. Genetically engineered peptides can associate with targeted transcription factors to block their function, both in vitro and in vivo[11]. An improved understanding of the recognition principles underlying protein–protein interactions is paving the path towards the rational design of protein–interface mimetics that are highly target-specific, potentially lowering the number of off-target effects in cells[2,6]. And, as Moellering et al. show, the ability to make relatively large agents that are cell-permeable, such as stapled helices, offers the tangible prospect of targeting the evasive regulators of gene networks that dictate cell fate and function. ■

Paramjit S. Arora is in the Department of Chemistry, New York University, New York 10003, USA. Aseem Z. Ansari is in the Department of Biochemistry and the Genome Center, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA.
e-mails: arora@nyu.edu;
ansari@biochem.wisc.edu

1. Moellering, R. E. et al. Nature **462,** 182–188 (2009).
2. Wells, J. A. & McClendon, C. L. Nature **450,** 1001–1009 (2007).
3. Jones, S. & Thornton, J. M. Proc. Natl Acad. Sci. USA **93,** 13–20 (1996).
4. Mapp, A. K. & Ansari, A. Z. ACS Chem. Biol. **2,** 62–75 (2007).
5. Schafmeister, C. E., Po, J. & Verdine, G. L. J. Am. Chem. Soc. **122,** 5891–5892 (2000).
6. Henchey, L. K., Jochim, A. L. & Arora, P. S. Curr. Opin. Chem. Biol. **12,** 692–697 (2008).
7. Wender, P. A. et al. Proc. Natl Acad. Sci. USA **97,** 13003–13008 (2000).
8. Walensky, L. D. et al. Science **305,** 1466–1470 (2004).
9. Darnell, J. E. Jr Nature Rev. Cancer **2,** 740–749 (2002).
10. Vazquez, A., Bond, E. E., Levine, A. J. & Bond, G. L. Nature Rev. Drug Discov. **7,** 979–987 (2008).
11. Oh, W. J., Rishi, V., Orosz, A., Gerdes, M. J. & Vinson, C. Cancer Res. **67,** 1867–1876 (2007).

# ARTICLES

# Predicting new molecular targets for known drugs

Michael J. Keiser[1,2]*, Vincent Setola[3]*, John J. Irwin[1], Christian Laggner[1], Atheir I. Abbas[4], Sandra J. Hufeisen[5], Niels H. Jensen[5], Michael B. Kuijer[3], Roberto C. Matos[3], Thuy B. Tran[3], Ryan Whaley[3], Richard A. Glennon[6], Jérôme Hert[1], Kelan L. H. Thomas[1,7], Douglas D. Edwards[1], Brian K. Shoichet[1] & Bryan L. Roth[3,5]

**Although drugs are intended to be selective, at least some bind to several physiological targets, explaining side effects and efficacy. Because many drug–target combinations exist, it would be useful to explore possible interactions computationally. Here we compared 3,665 US Food and Drug Administration (FDA)-approved and investigational drugs against hundreds of targets, defining each target by its ligands. Chemical similarities between drugs and ligand sets predicted thousands of unanticipated associations. Thirty were tested experimentally, including the antagonism of the $\beta_1$ receptor by the transporter inhibitor Prozac, the inhibition of the 5-hydroxytryptamine (5-HT) transporter by the ion channel drug Vadilex, and antagonism of the histamine $H_4$ receptor by the enzyme inhibitor Rescriptor. Overall, 23 new drug–target associations were confirmed, five of which were potent ($<$100 nM). The physiological relevance of one, the drug *N,N*-dimethyltryptamine (DMT) on serotonergic receptors, was confirmed in a knockout mouse. The chemical similarity approach is systematic and comprehensive, and may suggest side-effects and new indications for many drugs.**

The creation of target-specific 'magic bullets' has been a therapeutic goal since Ehrlich[1], and a pragmatic criterion in drug design for 30 years. Still, several lines of evidence suggest that drugs may have many physiological targets[2–5]. Psychiatric medications, for instance, notoriously act through multiple molecular targets, and this 'polypharmacology' is probably therapeutically essential[6]. Recent kinase drugs, such as Gleevec and Sutent, although perhaps designed for specificity, modulate several targets, and these 'off-target' activities may also be essential for efficacy[7,8]. Conversely, anti-Parkinsonian drugs such as Permax and Dostinex activate not only dopamine receptors but also 5-HT$_{2B}$ serotonin receptors, thereby causing valvular heart disease and severely restricting their use[9].

## Predicting drug polypharmacology

Drug polypharmacology has inspired efforts to predict and characterize drug–target associations[10–15]. Several groups have used phenotypic and chemical similarities among molecules to identify those with multiple targets[16,17], and early drug candidates are screened against molecular target panels[18]. To predict new targets for established drugs, a previous group looked for side-effects shared between two molecules[19], whereas another group linked targets by drugs that bind to more than one of them[20]. Indeed, using easily accessible associations, one can map 332 targets by the 290 drugs that bind to at least two of them, resulting in a network with 972 connections (Fig. 1a). It seemed interesting to calculate a related map that predicts new off-target effects.

Accordingly, we used a statistics-based chemoinformatics approach to predict new off-targets for 878 purchasable FDA-approved small-molecule drugs and 2,787 pharmaceutical compounds. Unlike bioinformatics methods, which might use the sequence or structural

similarity among targets, this similarity ensemble approach (SEA)[21] compares targets by the similarity of the ligands that bind to them, expressed as expectation values, adapting the BLAST algorithms[21–23] (other methods such as naive Bayesian classifiers[23,24] may also be used, see Supplementary Table 1). The approach thus captures ligand-based similarities among what would otherwise be considered disparate proteins. The 3,665 drugs were compared against 65,241 ligands organized into 246 targets drawn from the MDL Drug Data Report (MDDR) database[25], yielding 901,590 drug–target comparisons.

Most drugs had no significant similarities to most ligand sets. However, 6,928 pairs of drugs and ligand sets were similar, with expectation values (*E*-values) better than $1 \times 10^{-10}$. We analysed these predictions retrospectively against known associations and prospectively for unreported drug polypharmacology.

## Retrospective drug–target predictions

We first compared the predicted drug–target associations from the MDDR database against reported associations with affinities better than 1 μM in a second database, the World of Molecular Bioactivity (WOMBAT)[26]. For instance, the MDDR annotates Azopt (brinzolamide) only as an "antiglaucoma agent", but WOMBAT reports that it binds carbonic anhydrase II at 3 nM. Correspondingly, when screened internally against all MDDR molecular targets, SEA associated this drug with "carbonic anhydrase inhibitors" with an *E*-value of $8.3 \times 10^{-139}$. For 184 of the 746 drugs in WOMBAT, the predicted MDDR target agreed with the annotated WOMBAT target with *E*-values of $1 \times 10^{-10}$ or better, recapitulating 19% of the off-targets missing from the MDDR (Supplementary Table 2). Another 257 drug–target predictions were unannotated in either database, and may suggest new polypharmacology.

[1]Department of Pharmaceutical Chemistry, [2]Graduate Group in Bioinformatics, University of California San Francisco, 1700 4th Street, San Francisco, California 94143-2550, USA. [3]NIMH Psychoactive Drug Screening Program, Department of Pharmacology, [4]Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA. [5]Department of Pharmacology and Division of Medicinal Chemistry and Natural Products, The University of North Carolina Chapel Hill School of Medicine, Chapel Hill, North Carolina 27759, USA. [6]Department of Medicinal Chemistry, School of Pharmacy, Medical College of Virginia Campus, Virginia Commonwealth University, 410 North 12th Street, PO Box 980540, Richmond, Virginia 23298-0540, USA. [7]University of Michigan Health System, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109, USA.
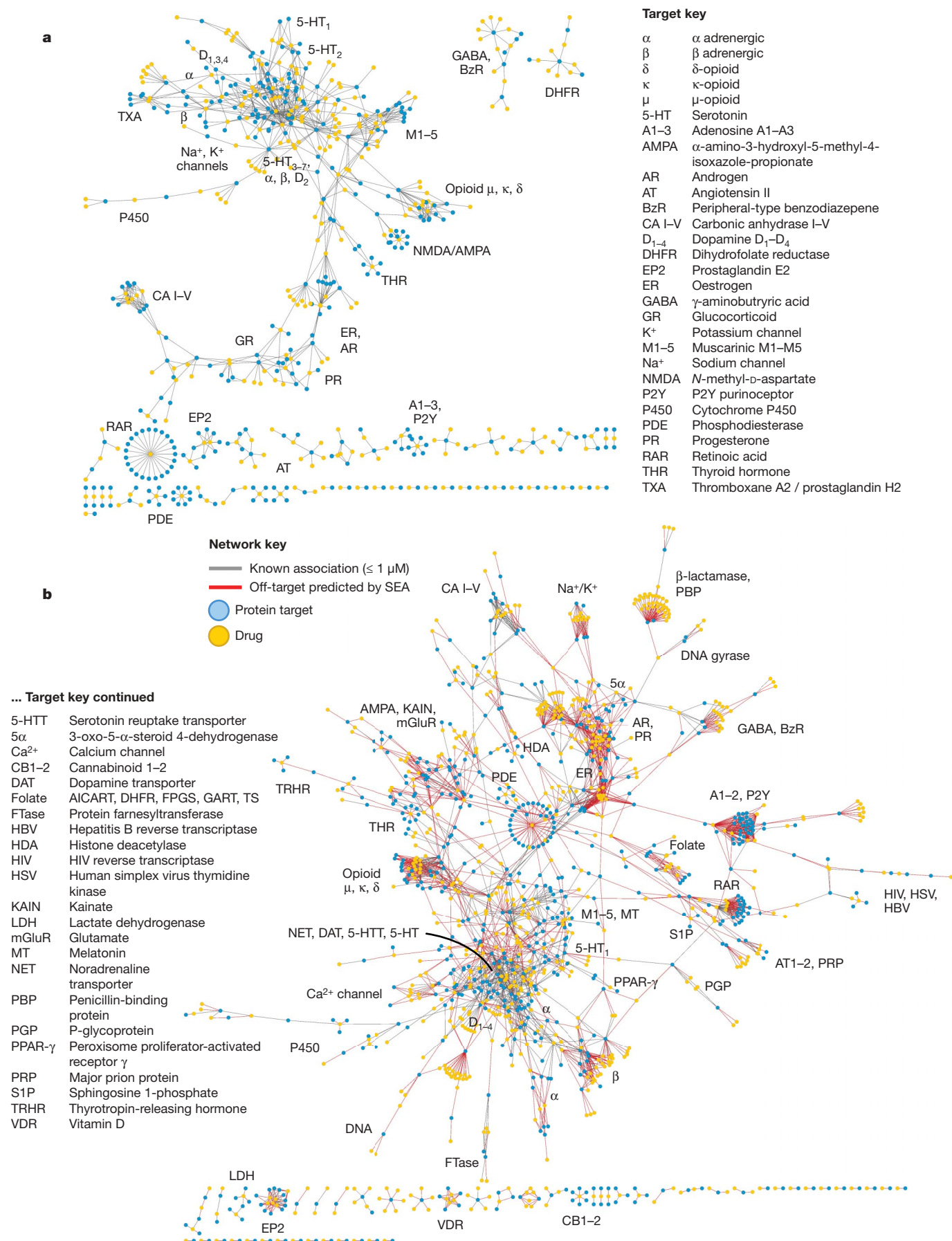*These authors contributed equally to this work.

**Target key**

| | |
|---|---|
| α | α adrenergic |
| β | β adrenergic |
| δ | δ-opioid |
| κ | κ-opioid |
| μ | μ-opioid |
| 5-HT | Serotonin |
| A1–3 | Adenosine A1–A3 |
| AMPA | α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate |
| AR | Androgen |
| AT | Angiotensin II |
| BzR | Peripheral-type benzodiazepene |
| CA I–V | Carbonic anhydrase I–V |
| $D_{1-4}$ | Dopamine $D_1$–$D_4$ |
| DHFR | Dihydrofolate reductase |
| EP2 | Prostaglandin E2 |
| ER | Oestrogen |
| GABA | γ-aminobutyric acid |
| GR | Glucocorticoid |
| $K^+$ | Potassium channel |
| M1–5 | Muscarinic M1–M5 |
| $Na^+$ | Sodium channel |
| NMDA | N-methyl-D-aspartate |
| P2Y | P2Y purinoceptor |
| P450 | Cytochrome P450 |
| PDE | Phosphodiesterase |
| PR | Progesterone |
| RAR | Retinoic acid |
| THR | Thyroid hormone |
| TXA | Thromboxane A2 / prostaglandin H2 |

**... Target key continued**

| | |
|---|---|
| 5-HTT | Serotonin reuptake transporter |
| 5α | 3-oxo-5-α-steroid 4-dehydrogenase |
| $Ca^{2+}$ | Calcium channel |
| CB1–2 | Cannabinoid 1–2 |
| DAT | Dopamine transporter |
| Folate | AICART, DHFR, FPGS, GART, TS |
| FTase | Protein farnesyltransferase |
| HBV | Hepatitis B reverse transcriptase |
| HDA | Histone deacetylase |
| HIV | HIV reverse transcriptase |
| HSV | Human simplex virus thymidine kinase |
| KAIN | Kainate |
| LDH | Lactate dehydrogenase |
| mGluR | Glutamate |
| MT | Melatonin |
| NET | Noradrenaline transporter |
| PBP | Penicillin-binding protein |
| PGP | P-glycoprotein |
| PPAR-γ | Peroxisome proliferator-activated receptor γ |
| PRP | Major prion protein |
| S1P | Sphingosine 1-phosphate |
| TRHR | Thyrotropin-releasing hormone |
| VDR | Vitamin D |

**Figure 1 | Drug–target networks, before and after predicting off-targets.** a, Known drug–target network. Each drug (gold) is linked to its known protein targets (cyan) by a grey edge. Each edge denotes a $K_i$ of 1 μM or better for that drug to its target. b, Predicted drug–target network. Drugs and proteins are linked as per the known drug–target network in a, but with the addition of red edges representing SEA off-target predictions with E-values $\leq 10^{-10}$.

A second retrospective test predicted targets for the 3,665 drugs uncharacterized in either database but known in the literature. Of the 6,928 drug off-targets predicted, we discarded 430 as highly similar by structure to known target ligands, and another 2,666 as trivial. This left 3,832 predictions, of which we inspected 184 by literature search and by interrogating other databases. Of these, 42 turned out to be known associations (Supplementary Table 3). For instance, when we screened the drug Revanil (lisuride) against the MDDR ligand–target sets, its best $E$-value was as an $\alpha_2$ adrenergic antagonist, and when we screened the drug Permax (pergolide) it had an $E$-value of $8.70 \times 10^{-29}$ as a $5\text{-HT}_{1D}$ receptor agonist. Consistent with these predictions, Revanil has been reported to bind adrenergic $\alpha_2$ at 0.055 nM and Permax binds the $5\text{-HT}_{1D}$ receptor at 13 nM (Supplementary Table 3), although neither activity was reported in the MDDR or WOMBAT databases.

### New drug–target predictions

For many of these 184 predictions we found no literature precedent. We therefore tested 30 predictions that were experimentally accessible to us. In radioligand competition binding assays, 23 of these (77%) yielded inhibition constants ($K_i$ values) less than 15 µM (lower $K_i$ values indicate higher affinity) (Tables 1 and 2 and Supplementary Fig. 1). Fifteen of these 23 were to aminergic G-protein-coupled receptors (GPCRs) (Table 1), and the remainder crossed major receptor classification boundaries (Table 2). For instance, the $\alpha_1$ antagonist Doralese was predicted and observed to bind to the dopamine $D_4$ receptor—both $\alpha_1$ and $D_4$ are aminergic GPCRs. Conversely, the HIV-1 reverse transcriptase (enzyme) inhibitor Rescriptor was predicted and observed to bind to histamine $H_4$ receptor (GPCR); this prediction crosses major target boundaries. For several predictions, we tested multiple receptor subtypes because the MDDR left these unspecified; for example, for a predicted '$\alpha_1$ adrenergic blocker', we tested the drug at $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ subtypes; we count these as a single target. In total, 14 drugs bound 23 previously unknown targets, with 13 having sub-micromolar and five having sub-100 nM affinities (Tables 1 and 2). In cases such as Doralese's, the affinity for the discovered off-target dopamine $D_4$, to which it binds with a $K_i$ of 18 nM (Fig. 2a), was better than that for its known therapeutic targets, $\alpha_{1A}$ and $\alpha_{1B}$ adrenergic receptors, for which its $K_i$ values are 611 and 226 nM, respectively[27].

The question arises as to how interesting and biologically relevant these new off-targets are. This can be evaluated by the following criteria: when the new targets contribute to the primary activity of the drug, when they may mediate drug side effects, or when they are unrelated by sequence, structure and function to the canonical targets. Although not all of the newly predicted off-targets fall into these three categories, several fall into each.

### New targets as primary sites of action

The new targets can improve our understanding of drug action. DMT is an endogenous metabolite and a notorious hallucinogen. Recently, the molecule was characterized as a $\sigma_1$-receptor regulator at micromolar concentrations, an association implicated in its hallucinogenic properties[28,29]. This surprised us because many drugs, including non-hallucinogens, bind promiscuously to the $\sigma_1$ receptor with higher affinity than DMT[30]. Also, the hallucinogenic characteristics of DMT are consistent with other hallucinogens thought to act through serotonergic receptors, some of which the molecule is known to bind[31–33]. We therefore screened DMT against the 1,133 WOMBAT targets. SEA predicted it to be similar against multiple serotonergic (5-HT) ligand sets, with $E$-values ranging from $9.2 \times 10^{-81}$ to $7.4 \times 10^{-6}$. Upon testing, we find DMT binds $5\text{-HT}_{1A}$, $5\text{-HT}_{1B}$, $5\text{-HT}_{1D}$, $5\text{-HT}_{2A}$, $5\text{-HT}_{2B}$, $5\text{-HT}_{2C}$, $5\text{-HT}_5A}$, $5\text{-HT}_6$ and $5\text{-HT}_7$ receptors with affinities ranging from 39 nM to 2.1 µM (Supplementary Table 4 and Supplementary Fig. 2). Of these, three were previously unknown (Table 1), and all had substantially greater affinities for DMT than that represented by its 14.75 µM dissociation constant

($K_d$) for $\sigma_1$ (ref. 28). To investigate further the role of serotonin receptors in DMT-induced hallucination, we turned to a cell-based assay and an animal model that are predictive of hallucinatory actions[34]. Consistent with SEA prediction, we find that DMT is not only a potent partial agonist at $5\text{-HT}_{2A}$ (Fig. 2g) as has been reported[31], but it also induces head twitch response in wild-type but not in $5\text{-HT}_{2A}$ knockout mice (Fig. 2h), which is new to this study. The half-maximum effective concentration ($EC_{50}$) of DMT at $5\text{-HT}_{2A}$ is 100-fold lower (better) than that observed for $\sigma_1$ (ref. 28). These observations support $5\text{-HT}_{2A}$ as the primary target for DMT's hallucinogenic effects.

Similarly, the new off-targets for Sedalande—a neuroleptic and anxiolytic derived from haloperidol—may illuminate this drug's therapeutic effects. Although used in psychiatric clinical trials as far back as the early 1960s[35], neither its mechanism of action in the central nervous system (CNS), nor that of the related Dimetholizine, is well understood. In addition to new activities against $\alpha_1$ adrenergic receptors (1.2–240 nM; Fig. 2b and Table 1), Dimetholizine was found to bind the $D_2$ and $5\text{-HT}_{1A}$ receptors, and Sedalande to bind the $5\text{-HT}_{1D}$ receptor (Table 1 and Supplementary Fig. 1). This probably contributes to the CNS activity of both drugs, given the association of the former with anxiety and aggression modulation, and the activity of many antipsychotics against the $D_2$ receptor. We also found analogues of Sedalande that were active against $5\text{-HT}_{1D}$, often at affinities comparable to or greater than those of Sedalande itself (Supplementary Table 5 and Supplementary Fig. 3). This supports the possibility of optimizing these drugs for new indications.
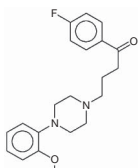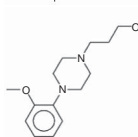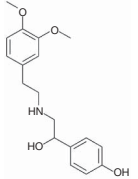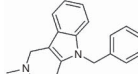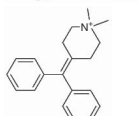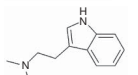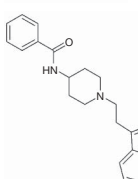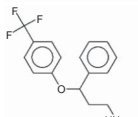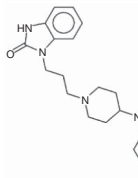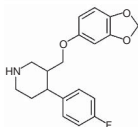
An example of a drug now being investigated for a new indication is Fabahistin. Used since the 1950s as a symptomatic antihistamine, Fabahistin is now being investigated for Alzheimer's disease. When screened against 1,133 WOMBAT targets, SEA found an extraordinary similarity to $5\text{-HT}_{5A}$ ligands, with an $E$-value of $2.0 \times 10^{-58}$. When we measured its binding to the $5\text{-HT}_{5A}$ receptor, Fabahistin had a $K_i$ of 130 nM (Fig. 2c and Table 1). This is another example of a drug whose new, off-target affinity is much better than that for its canonical $H_1$ receptor target[36]. Its activity against $5\text{-HT}_{5A}$ and related serotonergic receptors[37] may have implications for Fabahistin's role as an Alzheimer's disease therapeutic.

### Off-targets as side-effect mediators

Some of the new off-targets may contribute to a drug's adverse reactions. Motilium is an antiemetic and dopamine $D_{1/2}$ antagonist that achieves peak plasma concentrations of 2.8 µM[38] after intravenous administration. This formulation was withdrawn owing to adverse cardiovascular effects, with the FDA citing cardiac arrest, arrhythmias and sudden death[39]. Although Motilium binds the hERG channel with a half-maximum inhibitory concentration ($IC_{50}$) of 5 µM[40], the 71–710 nM affinities observed here against $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ may also contribute to these cardiovascular effects (Fig. 2d, Table 1 and Supplementary Fig. 1).

Similarly, the micromolar activity against the $\beta$ adrenergic receptors of the widely used selective serotonin reuptake inhibitor (SSRI) antidepressants Prozac and Paxil (Fig. 2e, Table 1 and Supplementary Fig. 1) may explain several of their adverse effects. The abrupt withdrawal of Paxil raises standing heart rate, a symptom of the SSRI discontinuation syndrome[41]. This is counterintuitive, as relieving blockade of serotonin reuptake should reduce synaptic serotonin, inconsistent with the cardiovascular syndrome[42]. $\beta$-blockade by these SSRIs may partially explain this effect because $\beta$-blockers induce a similar rebound tachycardia after abrupt withdrawal, owing to $\beta$ receptor upregulation and sensitization. Despite its higher affinity for $\beta$ receptors, Prozac has a longer half-life than Paxil, and its withdrawal does not induce SSRI discontinuation syndrome. Also, SSRIs and many $\beta$-blockers can induce sexual dysfunction[43]. Because both serotonergic and adrenergic signalling are involved in sexual response, the binding of Paxil and Prozac to the $\beta_1$-receptor may explain why they induce greater dysfunction than other SSRIs.

**Table 1 | Prediction and testing of new aminergic GPCR targets for drugs**

| Drug | | Pharmacological action | $E$-value | Predicted target | $K_i$ (nM) |
|------|---|------------------------|-----------|------------------|-----------|
| | Sedalande | Neuroleptic | $8.3 \times 10^{-136}$ | $\alpha_1$ adrenergic blocker* | $\alpha_{1A}$, 1.2; $\alpha_{1B}$, 14; $\alpha_{1D}$, 7.0 |
| | | | $1.7 \times 10^{-14}$ | 5-HT$_{1D}$ antagonist | 140 |
| | Dimetholizine | Antihistamine; antihypertensive | $1.6 \times 10^{-129}$ | $\alpha_1$ adrenergic blocker* | $\alpha_{1A}$, 70; $\alpha_{1B}$, 240; $\alpha_{1D}$, 170 |
| | | | $2.7 \times 10^{-113}$ | 5-HT$_{1A}$ antagonist | 110 |
| | | | $7.4 \times 10^{-56}$ | Dopamine D$_2$ antagonist | 180 |
| | Kalgut | Cardiotonic | $3.1 \times 10^{-79}$ | $\beta_3$ adrenergic agonist | $2.1 \times 10^3$ |
| | Fabahistin | Antihistamine | $5.7 \times 10^{-57}$ | 5-HT$_{5A}$ antagonist | 130 |
| | Prantal | Anticholinergic; antispasmodic | $5.5 \times 10^{-32}$ | $\delta$-opioid agonist | $1.4 \times 10^4$ |
| | $N,N$-dimethyltryptamine | Serotonergic hallucinogen | $3.1 \times 10^{-21}$ | 5-HT$_{1B}$ agonist | 130 |
| | | | $1.2 \times 10^{-13}$ | 5-HT$_{2A}$ agonist† | 130 |
| | | | $1.1 \times 10^{-7}$ | 5-HT$_{5A}$ antagonist | $2.1 \times 10^3$ |
| | | | $5.0 \times 10^{-6}$ | 5-HT$_7$ modulator | 210 |
| | Doralese | Adrenergic $\alpha_1$ blocker; antihypertensive; antimigraine | $2.8 \times 10^{-27}$ | Dopamine D$_4$ antagonist | 18 |
| | Prozac | 5-HT reuptake inhibitor; antidepressant | $3.9 \times 10^{-15}$ | $\beta$ adrenergic blocker* | $\beta_1$, $4.4 \times 10^3$ |
| | Motilium | Antiemetic; peristaltic stimulant | $4.8 \times 10^{-11}$ | $\alpha_1$ adrenergic blocker* | $\alpha_{1A}$, 71; $\alpha_{1B}$, 530; $\alpha_{1D}$, 710 |
| | Paxil | 5-HT reuptake inhibitor; antidepressant | $1.3 \times 10^{-7}$ | $\beta$ adrenergic blocker* | $\beta_1$, $1.0 \times 10^4$ |

$K_i$ values are accurate ±20% at two significant figures.
* For the targets marked, the reference data set did not specify the receptor subtype, requiring a separate assay for each one. For instance, the MDDR contains an '$\alpha_1$ adrenergic blocker' set, for which it was necessary to test the $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ subtypes.
† 5-HT$_{2A}$ is a known target of DMT, but is shown here with its retrospective SEA $E$-value for comparison purposes.
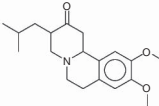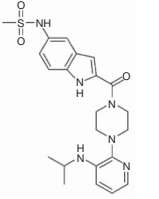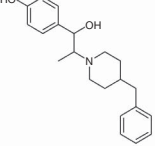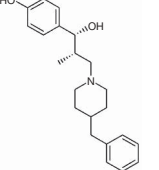
### Drug binding across major protein boundaries

Whereas many of the predicted off-targets occur among aminergic GPCRs, a target class for which cross-activity is well-known (see later)[44], four of the drugs bound to targets unrelated by sequence or structure to their canonical targets (Table 2). For instance, the reverse transcriptase (enzyme) inhibitor Rescriptor was predicted and shown to bind to the histamine H$_4$ receptor, a GPCR. These two targets share no evolutionary history, functional role, or structural similarity whatsoever. Intriguingly, although the $K_i$ value of

Rescriptor for the H$_4$ receptor is high at 5.3 µM (Table 2 and Supplementary Fig. 1), this is within its steady-state plasma concentration (minimum plasma concentration averages 15 µM) and is consistent with the painful rashes associated with Rescriptor use[45]; likewise, H$_4$ dysregulation has been associated with atopic dermatitis[46]. Similarly, the vesicular monoamine transporter (VMAT) inhibitor[47] Xenazine binds two different GPCRs at sub-micromolar concentrations (Table 2 and Supplementary Fig. 1). Despite its use over the last 50 years, Xenazine has not been reported to bind to any

**Table 2 | Prediction and testing of new cross-boundary targets for drugs**

| Drug | | Canonical target | E-value | Predicted target | $K_i$ (nM) |
|---|---|---|---|---|---|
|  | Xenazine | VMAT2 (transporter) | $1.4 \times 10^{-61}$ | $\alpha_2$ adrenergic receptor (GPCR) | $\alpha_{2A}$, 960; $\alpha_{2C}$, $1.3 \times 10^3$ |
|  | Rescriptor | HIV-1 reverse transcriptase (enzyme) | $1.1 \times 10^{-30}$ | Histamine $H_4$ receptor (GPCR) | $5.3 \times 10^3$ |
|  | Vadilex | NMDAR (ion channel) | $5.1 \times 10^{-13}$ | $\mu$-opioid receptor (GPCR) | $1.4 \times 10^3$ |
| | | | $2.0 \times 10^{-4}$ | 5-HTT; serotonin transporter (transporter) | 77 |
|  | RO-25-6981 | NMDAR (ion channel) | $1.5 \times 10^{-8}$ | 5-HTT; serotonin transporter (transporter) | $1.4 \times 10^3$ |
| | | | $1.9 \times 10^{-6}$ | Dopamine $D_4$ receptor (GPCR) | 120 |
| | | | $3.6 \times 10^{-6}$ | NET; noradrenaline transporter (transporter) | $1.3 \times 10^3$ |
| | | | $9.1 \times 10^{-5}$ | $\kappa$-opioid receptor (GPCR) | $3.1 \times 10^3$ |

$K_i$ values are accurate $\pm 20\%$ at two significant figures. The MDDR database did not specify the $\alpha_2$ adrenergic receptor subtype, requiring a separate assay for each one ($\alpha_{2A}$, $\alpha_{2C}$).

GPCR. Finally, the selective ion-channel inhibitors Vadilex and RO-25-6981 were predicted and found to bind to GPCRs and to transporters to which they were previously unknown to bind (Fig. 2f,



**Figure 2 | Testing new off-target activities. a–f,** Radioligand competition binding assays: Doralese at $D_4$ (**a**), Sedalande and Dimetholizine at $\alpha_{1D}$ (**b**), Fabahistin at 5-HT$_{5A}$ (**c**), Motilium at $\alpha_{1A}$ (**d**), Prozac at $\beta_1$ (**e**), and Vadilex (**f**) at the serotonin transporter. c.p.m., counts per minute; HEAT, 2-[$\beta$-(4-hydroxyphenyl)-ethyl-aminomethyl]tetralone; LSD, lysergic acid diethylamide. **g, h,** Investigating 5-HT$_{2A}$ as the target of DMT-induced hallucination: 5-HT$_{2A}$-mediated $Ca^{2+}$ response was measured after treating HEK 293 cells stably expressing the human 5-HT$_{2A}$ receptor with DMT (filled circles) or 5-HT (open circles) (**g**). The mean EC$_{50}$ $\pm$ s.e.m. of DMT was found to be $118 \pm 29$ nM (versus 5-HT's $6.6 \pm 0.4$ nM baseline, $n = 3$), with a maximum percentage relative (to 5-HT) efficacy (E$_{max}$) of $23 \pm 0.4\%$ ($n = 3$), confirming that DMT is a potent partial agonist at 5-HT$_{2A}$ receptors. DMT elicited head twitch behaviour only in 5-HT$_{2A}$ wild-type (WT) mice, confirming that it is a hallucinogenic 5-HT$_{2A}$ agonist (**h**). KO, knockout. **$P < 0.01$. In **a–f**, error bars represent the s.e.m. (of three replicates) from one representative experiment; in **h**, error bars represent the s.e.m. (of seven mice per genotype).

Table 2 and Supplementary Fig. 1). Although these ion-channel drugs have known polypharmacology (Fig. 3), a key point is that the new targets for these four drugs are unrelated to their main therapeutic targets except in the similarity of the ligands that modulate their activities.

More broadly, the protein target with the highest sequence similarity to any of a drug's known targets is rarely predicted by the SEA approach. Rather, the target predicted by ligand similarity is typically well down in the sequence-similarity ranking. Thus for Xenazine, the off-target $\alpha_2$ adrenergic receptor is the 78th most similar receptor to the known target VMAT2 and in fact has no significant similarity at all, with a PSI-BLAST E-value of 125 (Supplementary Table 6); and for Rescriptor, $H_4$ is the 167th most similar receptor to HIV-1 reverse transcriptase, and even for Prantal, the aminergic $\delta$-opioid receptor is only the 26th most similar receptor to its known muscarinic $M_3$ target.

Certain caveats merit mention. Not all of the new off-targets predicted here would surprise specialists. For instance, Dimetholizine has antihypertensive activity and so its affinity for adrenergic receptors is not wholly unanticipated. Similarly, Kalgut is classified as a selective $\beta_1$ agonist, thought to have little activity on other adrenergic receptors[48]. Whereas the observation that it does bind to the $\beta_3$

receptor goes against this classification, structurally this seems easy to credit (Table 1 and Supplementary Fig. 1). Indeed, 10 of the 14 drugs reported here are active against aminergic GPCRs (Fig. 3), and so their cross-activities against other aminergic GPCRs have some precedent[44]. Finally, although most of the drugs were active at their predicted off-targets, one-third were not; these are examples of the false-positives to which this method is susceptible (Supplementary Table 7). Thus, the anxiolytics Valium and Centrax scored well against cholecystokinin B ligands, the antipsychotic Emilace was predicted to bind 5-HT$_4$, the anaesthetic Duocaine the κ-opioid receptor, the antihypertensive Doralese neurokinin receptors, and the narcotic Dromoran and the bradycardic Zatebradine scored well against the D$_2$ and D$_1$ receptors. None of these bound their predicted off-targets with affinities better than 10 μM. SEA ignores pharmacophores in its predictions, comparing drugs to ligand sets based on all shared chemical patterns. This is at once a strength, in that it is model-free, and a weakness, in that it may predict activity for drugs that share many features with the ligands of a target, and yet miss a critical chemotype.

## Predicting polypharmacology on a large scale

Notwithstanding these caveats, it is the model-free nature of these predictions that allows a comprehensive exploration of drug–target interactions, most of which remain unexplored. We have focused on a thin slice of pharmacological targets, one dominated by aminergic drugs (Fig. 3). Stepping back to view the larger picture, 364 additional off-targets for 158 drugs are predicted with E-values better than

$1 \times 10^{-50}$, whereas 1,853 new off-targets are predicted with E-values better than $1 \times 10^{-10}$ (Fig. 1b). This compares to the only 972 off-target activities already annotated in the databases (Fig. 1a). The SEA and related chemoinformatics methods[16–20] provide tools to explore these associations systematically, both to understand drug effects and explore new opportunities for therapeutic intervention.

## METHODS SUMMARY

**Prediction of off-targets.** A collection of 3,665 FDA-approved and investigational drug structures was computationally screened against a panel of more than 1,400 protein targets. The drug collection was extracted from the MDL Comprehensive Medicinal Chemistry database. Each target was represented solely by its set of known ligands, which were extracted from three sources of annotated molecules: the MDL Drug Data Report, the WOMBAT[26], and the StARlite databases. The two-dimensional structural similarity of each drug to each target's ligand set was quantified as an E-value using the SEA[21].

**Experimental testing.** Predicted off-targets with strong SEA E-values were evaluated for novelty against orthogonal databases and the literature. Those off-targets without precedent were subjected to radioligand competition binding assays using standard techniques[49] at the NIMH Psychoactive Drug Screening Program. The role of 5-HT$_{2A}$ agonism in DMT-induced hallucination was examined in cell-based and in knockout mouse models[34]. Derivatives of Sedalande were identified in the ZINC[50] database by substructure search, and their affinities for 5-HT$_{1D}$ tested using standard techniques[49].

**Drug–target networks and out-group analysis.** Comprehensive networks of known drug–target associations (by WOMBAT) and predicted off-targets (by SEA) were constructed. Furthermore, SEA off-target predictions were compared to those derived from naive Bayesian classifiers and from PSI-BLAST[21–23] comparisons of a drug's known protein target(s) against the panel of potential protein targets.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.
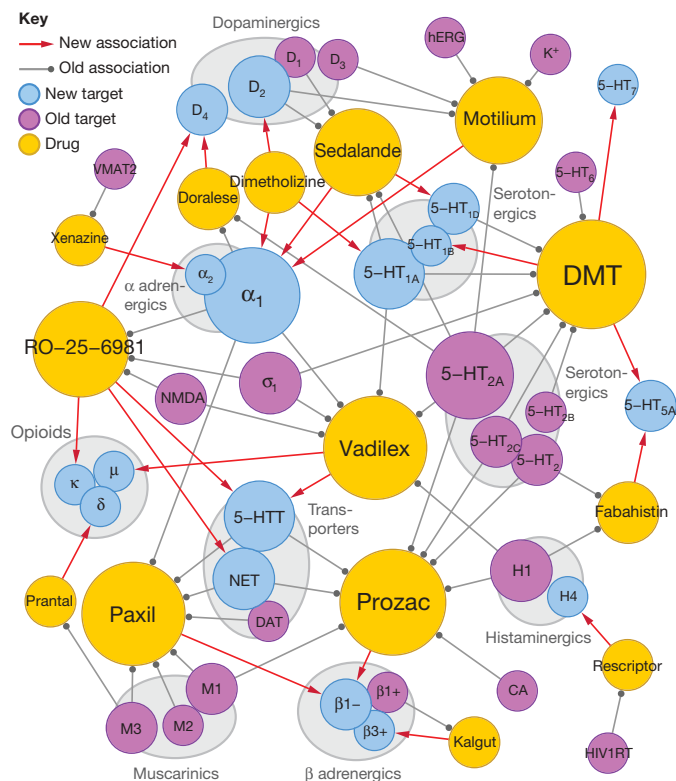
**Figure 3 | Discovered off-targets network.** Bipartite network where drugs (gold) are linked by grey edges to their known targets (violet) and by red arrows to their discovered off-targets (cyan). Grey edges denote binding at 1 μM or better, where these affinities are known. Node sizes increase with number of incident edges. Target abbreviations: 5-HT*x*, serotonin receptor type *x*; 5-HTT, serotonin transporter; β1+, β$_1$ adrenergic agonist; β1−, β$_1$ adrenergic antagonist; β3+, β$_3$ adrenergic agonist; σ$_1$, σ$_1$-receptor; CA, carbonic anhydrase; DAT, dopamine transporter; HIV1RT, HIV-1 reverse transcriptase; hERG, human Ether-a-go-go related gene channel; K$^+$, potassium channel; NET, noradrenaline transporter; NMDA, *N*-methyl-D-aspartate receptor; VMAT2, vesicular monoamine transporter 2.

1. Ehrlich, P. The theory and practice of chemotherapy. *Folia Serologica* 7, 697–714 (1911).
2. Peterson, R. T. Chemical biology and the limits of reductionism. *Nature Chem. Biol.* 4, 635–638 (2008).
3. Nobeli, I., Favia, A. D. & Thornton, J. M. Protein promiscuity and its implications for biotechnology. *Nature Biotechnol.* 27, 157–167 (2009).
4. Marona-Lewicka, D. & Nichols, D. E. Further evidence that the delayed temporal dopaminergic effects of LSD are mediated by a mechanism different than the first temporal phase of action. *Pharmacol. Biochem. Behav.* 87, 453–461 (2007).
5. Marona-Lewicka, D. & Nichols, D. E. WAY 100635 produces discriminative stimulus effects in rats mediated by dopamine D$_4$ receptor activation. *Behav. Pharmacol.* 20, 114–118 (2009).
6. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature Rev. Drug Discov.* 3, 353–359 (2004).
7. Rix, U. *et al.* Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. *Blood* 110, 4055–4063 (2007).
8. Hopkins, A. L. Network pharmacology. *Nature Biotechnol.* 25, 1110–1111 (2007).
9. Roth, B. L. Drugs and valvular heart disease. *N. Engl. J. Med.* 356, 6–9 (2007).
10. Bajorath, J. Computational analysis of ligand relationships within target families. *Curr. Opin. Chem. Biol.* 12, 352–358 (2008).
11. Oprea, T. I., Tropsha, A., Faulon, J. L. & Rintoul, M. D. Systems chemical biology. *Nature Chem. Biol.* 3, 447–450 (2007).
12. Newman, D. J. Natural products as leads to potential drugs: an old process or the new hope for drug discovery? *J. Med. Chem.* 51, 2589–2599 (2008).
13. Siegel, M. G. & Vieth, M. Drugs in other drugs: a new look at drugs as fragments. *Drug Discov. Today* 12, 71–79 (2007).
14. Miller, J. R. *et al.* A class of selective antibacterials derived from a protein kinase inhibitor pharmacophore. *Proc. Natl Acad. Sci. USA* 106, 1737–1742 (2009).
15. Walsh, C. T. & Fischbach, M. A. Repurposing libraries of eukaryotic protein kinase inhibitors for antibiotic discovery. *Proc. Natl Acad. Sci. USA* 106, 1689–1690 (2009).
16. Young, D. W. *et al.* Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nature Chem. Biol.* 4, 59–68 (2008).
17. Wagner, B. K. *et al.* Large-scale chemical dissection of mitochondrial function. *Nature Biotechnol.* 26, 343–351 (2008).
18. Krejsa, C. M. *et al.* Predicting ADME properties and side effects: the BioPrint approach. *Curr. Opin. Drug Discov. Dev.* 6, 470–480 (2003).
19. Campillos, M., Kuhn, M., Gavin, A. C., Jensen, L. J. & Bork, P. Drug target identification using side-effect similarity. *Science* 321, 263–266 (2008).
20. Paolini, G. V., Shapland, R. H. B., van Hoorn, W. P., Mason, J. S. & Hopkins, A. L. Global mapping of pharmacological space. *Nature Biotechnol.* 24, 805–815 (2006).

21. Keiser, M. J. *et al.* Relating protein pharmacology by ligand chemistry. *Nature Biotechnol.* **25**, 197–206 (2007).
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
23. Hert, J., Keiser, M. J., Irwin, J. J., Oprea, T. I. & Shoichet, B. K. Quantifying the relationships among drug classes. *J. Chem. Inf. Model.* **48**, 755–765 (2008).
24. Nigsch, F., Bender, A., Jenkins, J. L. & Mitchell, J. B. Ligand-target prediction using Winnow and naive Bayesian algorithms and the implications of overall performance statistics. *J. Chem. Inf. Model.* **48**, 2313–2325 (2008).
25. Schuffenhauer, A. *et al.* An ontology for pharmaceutical ligands and its application for in silico screening and library design. *J. Chem. Inf. Comput. Sci.* **42**, 947–955 (2002).
26. Olah, M. *et al.* in *Chemical Biology: From Small Molecules to Systems Biology and Drug Design* (eds Schreiber, S. L., Kapoor, T. M. & Wess, G.) 760–786 (Wiley-VCH, 2007).
27. Lomasney, J. W. *et al.* Molecular cloning and expression of the cDNA for the $\alpha_{1A}$-adrenergic receptor. The gene for which is located on human chromosome 5. *J. Biol. Chem.* **266**, 6365–6369 (1991).
28. Fontanilla, D. *et al.* The hallucinogen *N,N*-dimethyltryptamine (DMT) is an endogenous sigma-1 receptor regulator. *Science* **323**, 934–937 (2009).
29. Su, T. P., Hayashi, T. & Vaupel, D. B. When the endogenous hallucinogenic trace amine *N,N*-dimethyltryptamine meets the sigma-1 receptor. *Sci. Signal.* **2**, pe12 (2009).
30. Roth, B. L., Lopez, E., Patel, S. & Kroeze, W. K. The multiplicity of serotonin receptors: uselessly diverse molecules or an embarrasment of riches? *Neuroscientist* **6**, 252–262 (2000).
31. Smith, R. L., Canton, H., Barrett, R. J. & Sanders-Bush, E. Agonist properties of *N,N*-dimethyltryptamine at serotonin 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors. *Pharmacol. Biochem. Behav.* **61**, 323–330 (1998).
32. Kohen, R. *et al.* Cloning, characterization, and chromosomal localization of a human 5-HT$_6$ serotonin receptor. *J. Neurochem.* **66**, 47–56 (1996).
33. Pierce, P. A. & Peroutka, S. J. Hallucinogenic drug interactions with neurotransmitter receptor binding sites in human cortex. *Psychopharmacology (Berl.)* **97**, 118–122 (1989).
34. Abbas, A. I. *et al.* PSD-95 is essential for hallucinogen and atypical antipsychotic drug actions at serotonin receptors. *J. Neurosci.* **29**, 7124–7136 (2009).
35. Kurland, A. A., Mc, C. K. & Michaux, W. W. Clinical trial of haloanisone (R-2028) with hospitalized psychiatric patients. *J. New Drugs* **2**, 352–360 (1962).
36. Gankina, E. M. *et al.* Effect of some antihistamine preparations on binding of $^3$H-mepyramine and $^3$H-cimetidine to histamine receptors in rat brain. *Pharm. Chem. J* **26**, 373–376 (1992).
37. Gankina, E. M. *et al.* The effect of antihistaminic preparations on the binding of labelled mepyramine, ketanserin and quinuclidinyl benzilate in the rat brain [in Russian with English abstract]. *Eksp. Klin. Farmakol.* **56**, 22–24 (1993).
38. Heykants, J. *et al.* On the pharmacokinetics of domperidone in animals and man. IV. The pharmacokinetics of intravenous domperidone and its bioavailability in man following intramuscular, oral and rectal administration. *Eur. J. Drug Metab. Pharmacokinet.* **6**, 61–70 (1981).
39. FDA. Talk Paper: FDA Warns Against Women Using Unapproved Drug, Domperidone, to Increase Milk Production < http://www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/ucm173886.htm > (7 June 2004).
40. Stork, D. *et al.* State dependent dissociation of HERG channel inhibitors. *Br. J. Pharmacol.* **151**, 1368–1376 (2007).
41. Michelson, D. *et al.* Interruption of selective serotonin reuptake inhibitor treatment. Double-blind, placebo-controlled trial. *Br. J. Psychiatry* **176**, 363–368 (2000).
42. Berger, M., Gray, J. A. & Roth, B. L. The extended pharmacology of serotonin. *Annu. Rev. Med.* **60**, 355–366 (2009).
43. Waldinger, M. D., Hengeveld, M. W., Zwinderman, A. H. & Olivier, B. Effect of SSRI antidepressants on ejaculation: a double-blind, randomized, placebo-controlled study with fluoxetine, fluvoxamine, paroxetine, and sertraline. *J. Clin. Psychopharmacol.* **18**, 274–281 (1998).
44. Peters, J. U., Schnider, P., Mattei, P. & Kansy, M. Pharmacological promiscuity: dependence on compound properties and target specificity in a set of recent Roche compounds. *ChemMedChem* **4**, 680–686 (2009).
45. Scott, L. J. & Perry, C. M. Delavirdine: a review of its use in HIV infection. *Drugs* **60**, 1411–1444 (2000).
46. Dijkstra, D. *et al.* Human inflammatory dendritic epidermal cells express a functional histamine H4 receptor. *J. Invest. Dermatol.* **128**, 1696–1703 (2008).
47. Mehvar, R., Jamali, F., Watson, M. W. & Skelton, D. Pharmacokinetics of tetrabenazine and its major metabolite in man and rat. Bioavailability and dose dependency studies. *Drug Metab. Dispos.* **15**, 250–255 (1987).
48. Masanori, I., Tetsuya, T., Tomihiro, I., Taku, N. & Shigeyuki, T. β1-adrenergic selectivity of the new cardiotonic agent denopamine in its stimulating effects on adenylate cyclase. *Biochem. Pharmacol.* **36**, 1947–1954 (1987).
49. Jensen, N. H. *et al.* *N*-desalkylquetiapine, a potent norepinephrine reuptake inhibitor and partial 5-HT$_{1A}$ agonist, as a putative mediator of quetiapine's antidepressant activity. *Neuropsychopharmacology* **33**, 2303–2312 (2008).
50. Irwin, J. J. & Shoichet, B. K. ZINC–a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182 (2005).

**Author Contributions** B.K.S., J.J.I. and M.J.K. developed the ideas for SEA. M.J.K. wrote the SEA algorithms, undertook the calculations, and identified the off-targets reported here, typically vetted with J.J.I. and B.K.S., unless otherwise noted below. M.J.K. wrote the naive Bayesian classifier algorithms with assistance from J.H. With assistance from B.K.S. and J.J.I., C.L. identified off-targets for Fabahistin, K.L.H.T. identified off-targets for Prozac and Paxil, and D.D.E. identified the off-target for Rescriptor. V.S. and B.L.R. designed empirical tests of the predictions, analysed and interpreted data, and performed experiments. V.S., T.B.T., R.W., R.C.M., A.A., N.H.J. and M.B.K. performed empirical testing of the predictions. V.S., S.J.H. and R.A.G. generated materials for the experiments. M.J.K., B.K.S. and B.L.R. wrote the manuscript with contributions and review from V.S. All authors discussed the results and commented on the manuscript.

## METHODS

**Ligand sets.** We extracted ligand sets from databases that annotate molecules by therapeutic or biological category. For example, the 2006.1 MDDR contains 518 molecules annotated as $\alpha_1$ adrenergic receptor blockers, which we grouped into a single '$\alpha_1$ adrenergic blocker' set.

As ligand reference sources, we used the three databases shown in Supplementary Table 8. The first was a subset of the 2006.1 MDDR, prepared as previously described[21,23]. The second was the 2006.2 WOMBAT database[26], processed as earlier. We collapsed WOMBAT targets across species and organized them into inhibitory, activating and unspecified-binding classes. All ligands with affinities worse than 1 μM to their targets were removed. This left 1,133 classes built from 191,943 ligands with median and mean of 37 and 169 ligands per target class. The third database was StARlite, which we also processed as above. We extracted StARlite annotations at the two highest confidence levels (5 and 7), discarded those with affinities worse than 1 μM, and organized them into target classes. This yielded 1,158 classes built from 111,329 ligands with a median and mean of 43 and 186 ligands per target class.

For drugs and bioactive molecules, we used the 2004 MDL Comprehensive Medicinal Chemistry database (CMC) of 7,517 compounds. Drugs were processed identically to the ligands above. Filtering by vendor availability (as reported in MDL 2006.3 Available Chemical Directory (ACD), MDL 2006.1 Screening Compounds Directory, and ZINC[50]) yielded 3,665 unique purchasable drugs.

The 1,216 drugs used to link protein targets in the drug-target networks (Fig. 1) were downloaded from the 2008 EPA Distributed Structure-Searchable Toxicity (DSSTox) Database at http://www.epa.gov/NCCT/dsstox/, and prepared as earlier.

**Ligand activity predictions.** We compared each drug individually against each set of ligands. Molecules were represented by two topological descriptors: 2048-bit Daylight[51] and 1024-bit folded ECFP_4 fingerprints[23]. We used the SEA[21,23] using each descriptor separately and chose top-scoring hits (that is, those with small *E*-values) from each such 'screen' independently.

The initial SEA screens of 3,665 CMC drugs against 246 MDDR targets yielded 901,590 drug–target comparisons, and we subjected these to retrospective literature analysis and prospective empirical testing. However, we later extended SEA screens to WOMBAT and StARlite databases, comprising 4,152,445 and 4,244,070 drug–target comparisons, respectively. We have not mined these expanded SEA screens for retrospective validations; instead we used them only to conduct prospective tests. Supplementary Table 8 records the screen (that is, database) from which each prediction in Tables 1 and 2 is derived.

To compare SEA predictions against those of naive Bayesian classifiers (Supplementary Table 1), we implemented a Laplacian-corrected naive Bayesian classifier with Avidon weighting, as previously described[23].

**Drug–target and target–target networks.** The drug–target networks in Fig. 1 are bipartite; along any given path, nodes alternate between protein targets and the drugs that link them. Targets are from WOMBAT and drugs are from EPA DSSTox. Red edges denote SEA predictions with *E*-values $\leq 10^{-10}$. Predictions already reported in WOMBAT at $K_i \leq 1$ μM are shown as additional grey edges. All networks were generated in Cytoscape 2.6.1 (ref. 52).

Figure 3 is a bipartite graph linking drugs from Tables 1 and 2 with protein targets. Grey edges link drugs to known targets from manual literature and database search. Grey edges denote binding at $\leq 1$ μM, except when no $K_i$ value was available; in these three cases (Xenazine–VMAT2, Prantal–$M_3$ and Fabahistin–$H_1$), the link was included for completeness.

**WOMBAT out-group analysis.** We mapped 204 MDDR activity classes to matching WOMBAT targets in two phases. In the first, we mapped 87 MDDR activity classes using EC numbers from the Schuffenhauer ontology[25] to those present in WOMBAT. We second mapped a further 118 non-enzyme MDDR activity classes by supervised sub-phrase matching (Supplementary Tables 9 and 10). Although this mapping is not guaranteed to be exhaustive, it is correct to the best of our knowledge.

We then extracted all molecules marked 'drug' in WOMBAT (746 unique). Using SEA, we compared them against the mapped MDDR classes only, and discarded all trivial hits (that is, those in which the drug was already annotated in that MDDR class as a ligand). We asked how many of these were, in retrospect, substantiated by the existing WOMBAT annotations, at affinities $\leq 1$ μM.

**Sequence similarity comparison.** We associated each drug in Fig. 3 with the human FASTA sequences of its known and its new protein targets, using http://www.uniprot.org. We ran these sequences by PSI-BLAST (BLAST version 2.2.14, default parameters)[21–23] against a subset of MDDR targets that we prepared as previously described[21]. For each SEA prediction in Fig. 3, we reported the best direct PSI-BLAST match (along with its *E*-value and ranking) of the new protein target to any of that drug's previously known protein targets (Supplementary Table 6). Our goal was to address whether starting with the best choice from a drug's known protein targets, we could recapitulate each SEA prediction solely by sequence similarity.

**Experimental testing.** Radioligand binding and functional assays were performed as previously described[49,53]. Detailed experimental protocols are available on the NIMH PDSP website at http://pdsp.med.unc.edu/UNC-CH%20Protocol%20Book.pdf.

**Mice.** All experiments were approved by the Institutional Animal Care and Use Committee at the University of North Carolina, Chapel Hill. Mice were housed under standard conditions: 12 h light/dark cycle and food and water ad libitum.

**Head twitch.** Littermate pairs of 5-HT$_{2A}$ wild-type and knockout mice were pretreated for 2 h with 75 mg kg$^{-1}$ pargyline, intraperitoneally, prepared in sterile saline (0.9% NaCl) (Sigma-Aldrich). Mice were then injected with sterile saline or 1.0 mg kg$^{-1}$ DMT, intraperitoneally, prepared in sterile saline and moved to a new cage. Head twitch behaviour, which consists of a rapid, rotational flick of the head about the axis of the neck, was counted over 15 min. We have determined that trained observers count the same number of head twitches whether blinded or unblinded to genotype (data not shown). We confirmed that this was the case with three littermate pairs, and the rest of the studies were performed by one unblinded observer[34].

51. James, C., Weininger, D. & Delany, J. *Daylight Theory Manual* (Daylight Chemical Information Systems Inc., 1992–2005).
52. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
53. Roth, B. L. *et al.* Salvinorin A: a potent naturally occurring nonnitrogenous κ opioid selective agonist. *Proc. Natl Acad. Sci. USA* **99**, 11934–11939 (2002).

nature

# ARTICLES

# Direct inhibition of the NOTCH transcription factor complex

Raymond E. Moellering[1,2,3], Melanie Cornejo[4], Tina N. Davis[6], Cristina Del Bianco[5], Jon C. Aster[5], Stephen C. Blacklow[5], Andrew L. Kung[6], D. Gary Gilliland[4,7], Gregory L. Verdine[1,3] & James E. Bradner[2,3,8]

**Direct inhibition of transcription factor complexes remains a central challenge in the discipline of ligand discovery. In general, these proteins lack surface involutions suitable for high-affinity binding by small molecules. Here we report the design of synthetic, cell-permeable, stabilized α-helical peptides that target a critical protein–protein interface in the NOTCH transactivation complex. We demonstrate that direct, high-affinity binding of the hydrocarbon-stapled peptide SAHM1 prevents assembly of the active transcriptional complex. Inappropriate NOTCH activation is directly implicated in the pathogenesis of several disease states, including T-cell acute lymphoblastic leukaemia (T-ALL). The treatment of leukaemic cells with SAHM1 results in genome-wide suppression of NOTCH-activated genes. Direct antagonism of the NOTCH transcriptional program causes potent, NOTCH-specific anti-proliferative effects in cultured cells and in a mouse model of NOTCH1-driven T-ALL.**

Transcription factors are master regulators of cell state. Commonly activated by genetic events or upstream signalling pathways, they mediate the neoplastic phenotype and maintain tissue specification in cancer. As such, they are highly desirable targets for ligand discovery[1]. Owing to expansive protein–protein interfaces and a general absence of hydrophobic pockets, transcription factors have proven among the most chemically intractable of all therapeutic targets. With the exception of nuclear hormone receptors that have evolved the ability to bind natural small-molecule ligands, potent and specific inhibitors of human transcription factors have not been realized. Here we report the successful development of a direct-acting antagonist of an oncogenic transcription factor, NOTCH1.

NOTCH proteins participate in conserved pathways that regulate cellular differentiation, proliferation and death[2,3]. Mammalian NOTCH receptors (NOTCH1–4 in humans) are single-pass transmembrane proteins that transmit juxtacrine signals initiated by ligands of the Delta, Serrate or Lag-2 family. Ligand binding to the extracellular domain of NOTCH1 initiates sequential proteolytic processing events catalysed respectively by an ADAM family metalloprotease and a γ-secretase complex, resulting in cytoplasmic release of the intracellular domain of NOTCH1 (ICN1)[4–6]. ICN1 then translocates to the nucleus and loads onto the DNA-bound transcription factor CSL[7]. The engagement of ICN1 with CSL creates a long, shallow groove along the interface of the two proteins that serves as a binding surface for co-activator proteins of the mastermind-like (MAML) family[8,9]. The resulting ICN–CSL–MAML ternary complex then recruits the core transcription machinery, effecting activation of NOTCH-dependent target genes.

The duration and strength of NOTCH signalling is normally tightly controlled. Whereas loss-of-function mutations have been observed in a variety of diseases[10–12], gain-of-function mutations in the NOTCH pathway are causally linked with cancer. Indeed, human NOTCH1 was first discovered owing to its involvement in a t(7;9) chromosomal translocation observed in patients with T-ALL[13].

Subsequently, various activating mutations in NOTCH1 have been discovered in >50% of patients with T-ALL[14]. Recently, further aberrations that potentiate NOTCH signalling have been identified, including loss-of-function of the NOTCH1 E3 ubiquitin ligase FBXW7 and the intracellular NOTCH inhibitor NUMB, in T-ALL and breast adenocarcinoma, respectively[15,16]. Activated NOTCH signalling has also been observed in cancers of the lungs, ovaries, pancreas and gastrointestinal tract as well as in melanoma, multiple myeloma and medulloblastoma[17–23].

Efforts to antagonize the NOTCH pathway have relied on blocking the generation of ICN using small-molecule inhibitors of the γ-secretase complex (GSIs)[24,25]. These molecules are not strictly NOTCH-specific, as they indiscriminately block the many signalling pathways downstream of γ-secretase[26]. T-ALL patients treated with GSIs suffer dose-limiting gastrointestinal toxicity, the origin of which is uncertain but it may result from chronic blockade of NOTCH1 and NOTCH2 processing[27]. Some cell lines containing activating NOTCH1 mutations are resistant to GSIs, and those that do respond commonly undergo growth arrest rather than apoptosis. These observations underscore the mechanistic utility and potential therapeutic value of NOTCH antagonists that act by directly targeting the NOTCH transactivation complex.

## Stapled α-helical peptides targeting the NOTCH complex

A dominant-negative fragment of MAML1 (residues 13–74; termed dnMAML1) has been shown to antagonize NOTCH signalling and cell proliferation when expressed in T-ALL cell lines[28,29]. In X-ray structures of the *Homo sapiens*[30] and *Caenorhabditis elegans*[31] core ICN–CSL–MAML complexes, this dnMAML1 polypeptide forms a nearly continuous α-helix that engages an elongated groove formed by the assembly of ICN1 and CSL (Fig. 1a). This α-helical interaction motif suggested that the NOTCH transactivation complex might be suitable for targeting by helix-mimetics such as hydrocarbon-stapled α-helical peptides, which in other systems have proven capable of targeting intracellular protein–protein interactions[32,33]. Specifically, we reasoned
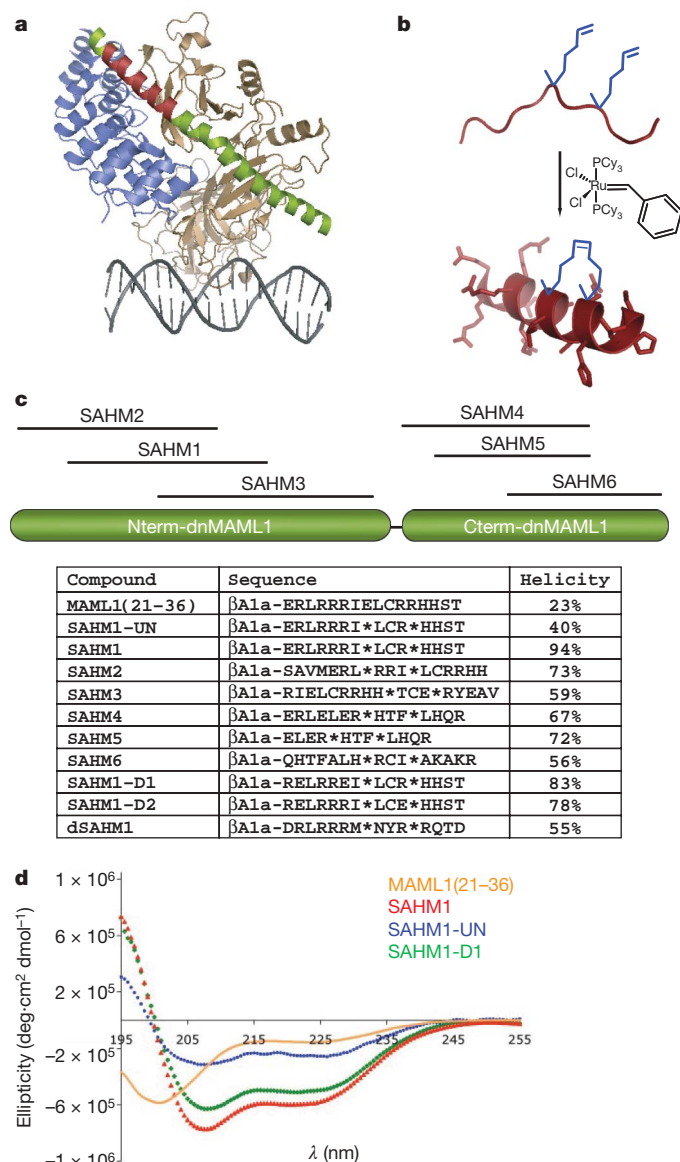
**Figure 1 | Design of MAML1-derived stapled peptides targeting NOTCH1–CSL. a**, Structure of the NOTCH1 ternary complex (Protein Data Bank (PDB) accession 2F8X): CSL (tan), DNA (grey), dnMAML1 (green) and ICN1 (blue). The 16-amino-acid stretch of MAML1 targeting ICN1 and CSL is shown in red, and was used to design the stapled peptide SAHM1. **b**, Schematic of peptide stapling. A non-natural alkenyl amino acid ($S_5$) is incorporated at two positions in the peptide chain and then cross-linked by ring-closing olefin metathesis. **c**, Schematic, sequences and helical character of MAML1-derived SAHM peptides. βAla denotes a β-alanine spacer. Asterisks denote the location of $S_5$ residues, which are cross-linked in all SAHM peptides except SAHM1-UN. **d**, Circular dichroism spectroscopy of four MAML1 derived peptides illustrating the incremental effects of synthetic modification.

| Compound | Sequence | Helicity |
|---|---|---|
| MAML1(21–36) | βA1a-ERLRRRIELCRRHHST | 23% |
| SAHM1-UN | βA1a-ERLRRRI*LCR*HHST | 40% |
| SAHM1 | βA1a-ERLRRRI*LCR*HHST | 94% |
| SAHM2 | βA1a-SAVMERL*RRI*LCRRHH | 73% |
| SAHM3 | βA1a-RIELCRRHH*TCE*RYEAV | 59% |
| SAHM4 | βA1a-ERLELER*HTF*LHQR | 67% |
| SAHM5 | βA1a-ELER*HTF*LHQR | 72% |
| SAHM6 | βA1a-QHTFALH*RCI*AKAKR | 56% |
| SAHM1-D1 | βA1a-RELRREI*LCR*HHST | 83% |
| SAHM1-D2 | βA1a-RELRRRI*LCE*HHST | 78% |
| dSAHM1 | βA1a-DRLRRRM*NYR*RQTD | 55% |

that a stapled fragment of dnMAML1 might prevent binding of full-length MAML1 to the ICN1–CSL complex, thereby depriving the complex of its transcriptional activation function despite the presence of an upstream NOTCH signal. Relative to unmodified peptides, hydrocarbon-stapled peptides have shown improved binding affinity, metabolic stability and serum half-life. Operationally, peptide stapling is accomplished by co-synthetic incorporation of a non-natural amino acid, $S_5$, at neighbouring positions along one face of the α-helix (*i* and *i* + 4 positions), followed by ring-closing olefin metathesis (Fig. 1b).

The structure of the human NOTCH1 ternary complex was used as the basis for the design of a series of stapled α-helical peptides derived

from MAML1 (SAHMs). Six candidate peptides were designed that together scan the entire contact surface of dnMAML1 with ICN1–CSL (Fig. 1c). Functional evaluation of this panel led to the selection of a stapled peptide designated SAHM1, which spans residues Glu 21 to Thr 36. As shown in Fig. 1d and Supplementary Fig. 1, stapling conferred upon SAHM1 a marked increase in helical character, as compared with its unmodified and modified but unstapled counterparts (MAML1(21–36) and SAHM1-UN, respectively). As specificity controls for functional studies[34], we generated mutant peptides SAHM1-D1 and SAHM1-D2, as well as *d*SAHM1, the corresponding peptide from the *Drosophila* Mastermind (MAM) protein.

Analysis of cell penetration using fluorescein isothiocyanate (FITC)-labelled peptides and quantitative epifluorescence microscopy showed robust and roughly equivalent cellular uptake by FITC–SAHM1, FITC–SAHM2, FITC–SAHM3, FITC–SAHM6 and the negative control peptides FITC–SAHM-D1, FITC–SAHM-D2 and FITC–*d*SAHM1. Cell penetration was less efficient for FITC–SAHM4 and negligible for FITC–-SAHM5 (Supplementary Fig. 2a, b). The uptake of FITC–SAHM1 was diminished at 4 °C relative to 37 °C, and was unaffected by rhodamine conjugation, consistent with an active, endocytic peptide import mechanism as observed previously with stapled peptides (Supplementary Fig. 2c, d). The molecules in this study are considerably smaller than the ~25-kilodalton (kDa) exclusion size of the nuclear pore complex, and thus as expected, FITC–SAHM1 was distributed throughout the cytoplasm and nucleus (Supplementary Fig. 2e).

**SAHMs bind the ICN1–CSL complex competitively with MAML1**

SAHM binding to the NOTCH complex was first investigated using an *in vitro* pull-down assay. In brief, a glutathione-*S*-transferase (GST)-tagged fragment of ICN1 bearing only the CSL-binding domain (GST–RAMANK) was immobilized on glutathione-labelled agarose beads and used to precipitate CSL and dnMAML1 in the presence and absence of SAHM1 (Fig. 2a). SAHM1 inhibited dnMAML1 binding competitively, but did not perturb binding of CSL to RAMANK.

SAHM binding was next measured quantitatively using fluorescence polarization spectroscopy. Soluble FITC–SAHM1 was found to bind the pre-formed RAMANK–CSL complex with a dissociation constant ($K_d$) = 0.12 ± 0.02 μM, whereas the unmodified FITC–MAML1(21–36) peptide bound the complex with markedly diminished affinity (Fig. 2b). Thus, helix stabilization promotes target binding. The mutant peptides FITC–SAHM1-D1, FITC–SAHM1-D2 and FITC–*d*SAHM1 bound RAMANK–CSL considerably less avidly than FITC–SAHM1, prompting the use of these peptides as negative controls in subsequent functional studies. Notably, fluorescence polarization of FITC–SAHM1 was decreased by unlabelled competitor dnMAML1, confirming the overlap between the SAHM1 and the dnMAML binding sites on RAMANK–CSL (Fig. 2c).

Kinetic insights into complex assembly and SAHM binding were gained using surface plasmon resonance (SPR). A constant amount of GST–RAMANK was immobilized on an immunoglobulin (anti-GST) biosensor surface and exposed to increasing concentrations of CSL. The resulting dose-dependent response data were best fit to a two-step association with rapid on- and off-rates and overall micromolar affinity (Fig. 2d), and are consistent with a published model supporting a two-step association[34]. Next, SPR was used to analyse binding of SAHMs to the pre-assembled RAMANK–CSL complex. A SAHM1 derivative containing an amino-terminal biotin moiety (bioSAHM1) was immobilized on a streptavidin sensor surface. Pre-incubated, equimolar RAMANK and CSL were introduced in increasing concentrations. Curve-fitting of these kinetic data indicated that bioSAHM1 bound RAMANK–CSL with an apparent $K_d$ closely matching the value observed by fluorescence polarization (Fig. 2e). Again, SAHM1-D1 showed substantially decreased affinity (Fig. 2f).
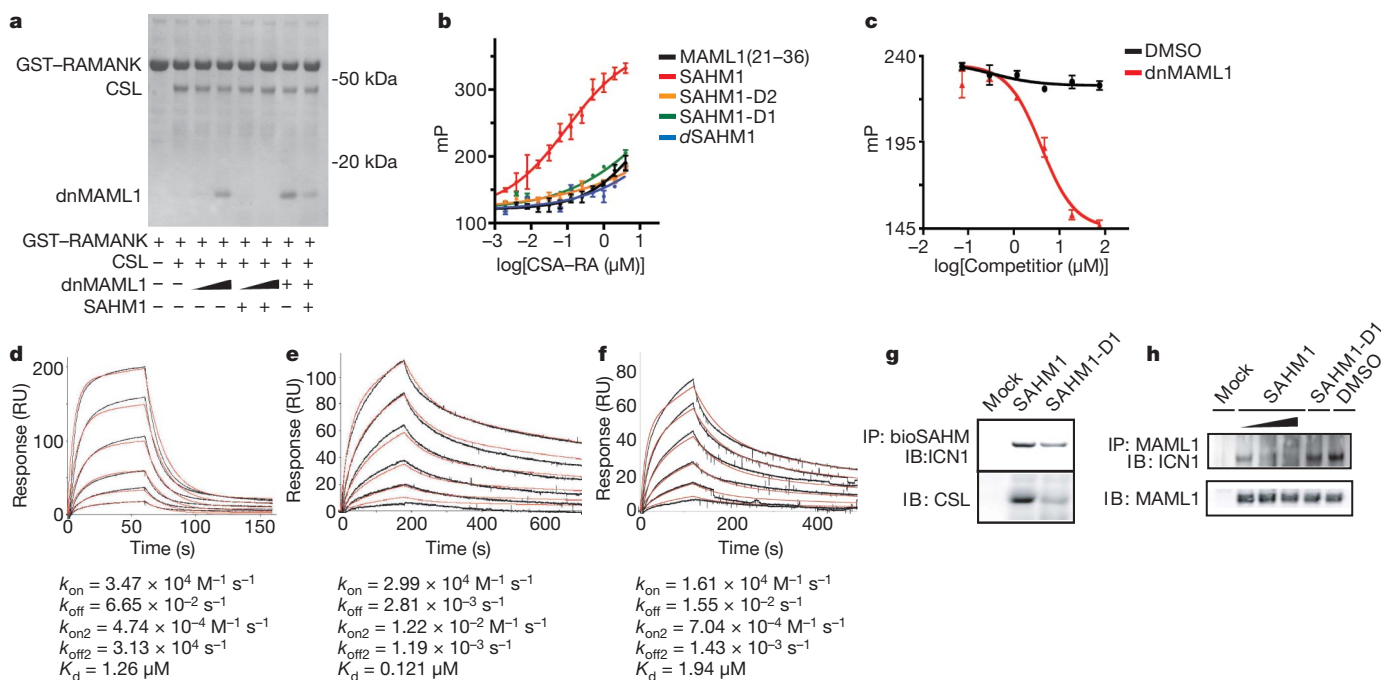
183

**Figure 2 | SAHM1 specifically engages the NOTCH1 transactivation complex. a**, *In vitro* assembly of the NOTCH1 complex. Bead-immobilized RAMANK protein was incubated as indicated with CSL (0.5 µM), dnMAML1 (0.5 µM for lanes 3 and 5 (from left); 2.5 µM for lanes 4 and 6; 5 µM for lanes 7 and 8), and SAHM1 (10 µM). Bound proteins were washed, eluted and resolved by gel electrophoresis (Coomassie). **b**, Fluorescence polarization of FITC–SAHM peptides binding to RAMANK–CSL. **c**, Direct competition between unlabelled dnMAML1 and FITC–SAHM1. Concentrations of FITC–SAHM1 (15 nM) and RAMANK–CSL (0.6 µM) were held constant. dnMAML1 $IC_{50} = 3.9 \pm 0.9$ µM. **d**, CSL binding to immobilized RAMANK by SPR. Black curves represent sensogram data and the red curve denotes fit to a two-step kinetic model. Binding constants are shown. $k_{on}$, association rate; $k_{off}$, dissociation rate; RU, response units. **e, f**, Binding of RAMANK–CSL complexes to immobilized bioSAHM1 (**e**) and bioSAHM1-D1 (**f**). **g**, bioSAHM1 and bioSAHM1-D1 pull-down assays in KOPT-K1 lysates. Bound protein fractions were probed with antibodies specific for ICN1 (top) and CSL (bottom). **h**, Competitive co-immunoprecipitation of endogenous ICN1 by MAML1 in the presence of vehicle, SAHM1 (0.5, 1 and 10 µM from left to right) or SAHM1-D1 (10 µM). Unless noted otherwise data represent the mean $\pm$ s.d. ($n = 3$).

To determine whether SAHMs bind endogenous NOTCH proteins present in human T-ALL cells, we performed pull-down assays in KOPT-K1 cellular lysates. BioSAHM1 beads were found to pull down full-length ICN1 and CSL, whereas bioSAHM1-D1 was markedly less avid (Fig. 2g). Next we examined the ability of SAHMs to compete with endogenous MAML1, again in KOPT-K1 cellular lysates. After immunoprecipitation of MAML1, the presence of ICN1 was detected by immunoblot. This association was diminished in a dose-dependent manner by SAHM1 but not SAHM-1D (Fig. 2h). Together, these experiments demonstrate conclusively that SAHM1 binds ICN1 and CSL and thereby directly antagonizes recruitment of the MAML1 co-activator to the ICN1–CSL complex.

## SAHM1 specifically represses NOTCH1 target gene expression

To explore the effects of SAHMs on transcriptional activity, we first used an established reporter-gene assay in which firefly luciferase is transcriptionally regulated by constitutively activated *NOTCH1* (ref. 35). SAHM1 treatment resulted in near complete repression of luciferase activity, comparable to a known GSI (DAPT)[25]. SAHM2 conferred a twofold signal reduction. SAHM4, SAHM5, SAHM6, SAHM1-D1, SAHM1-D2 and *d*SAHM1 were inactive (Fig. 3a). Reporter gene repression by SAHM1 was observed to be dose-dependent, with a half-maximum inhibitory concentration ($IC_{50}$) of $6.5 \pm 1.6$ µM (Fig. 3b). In a second reporter-gene assay monitoring β-lactamase transcriptionally regulated by ICN1 (ref. 36), SAHM1 again exhibited dose-dependent repression compared to SAHM1-D1 or vehicle alone (Supplementary Fig. 3a).

We next studied the effect of SAHMs on the expression of NOTCH target genes in NOTCH1-dependent T-ALL, quantified by PCR with reverse transcription (RT–PCR). A human T-ALL cell line (KOPT-K1) containing activating mutations in the heterodimerization (HD)

and degradation domains (PEST) of *NOTCH1* was treated with SAHM1, SAHM1-D1 and vehicle. Decreased expression of the NOTCH1 target genes *HES1*, *MYC* and *DTX1* was uniquely observed after treatment with SAHM1 (Fig. 3c). A consistent repressive effect of SAHM1 on NOTCH1-dependent gene expression was observed across a panel of human T-ALL cell lines (Fig. 3d), containing diverse mutant *NOTCH1* alleles (Supplementary Table 1). SAHM1 had no effect on ICN1 protein stability compared to DAPT (Supplementary Fig. 3b), supporting transcriptional inhibition downstream of ICN1 production.

## SAHM1 triggers global suppression of NOTCH1 signalling

Transcriptional profiling has recently been demonstrated to establish mechanistic connectivity between dissimilar compounds acting on common targets or target pathways[37]. Gene set enrichment analysis (GSEA) has also emerged as a robust method for comparing expression profiles corresponding to defined cellular states[38]. Integrating these two methodologies, we performed gene expression profiling as multidimensional phenotyping to first measure the global transcriptional effects of SAHM1 and then compare these effects with those produced by GSI treatment. First, triplicate data sets were generated from SAHM1 and vehicle-treated KOPT-K1 and HPB-ALL cells and analysed on Affymetrix oligonucleotide microarrays. Supervised hierarchical clustering and rank-ordering identified numerous canonical NOTCH1 target genes including *HES4*, *DTX1*, *HES1* and *MYC* among the top downregulated genes by SAHM1 (Fig. 3e). Next, a set of transcripts downregulated by GSI in T-ALL cell lines was curated from a published expression profile (referred to here as the GSI-NOTCH gene set)[39]. Enrichment for this gene set within the SAHM1 expression profile was studied by GSEA, revealing a strong, statistically significant correlation (enrichment score (ES) = −0.89,
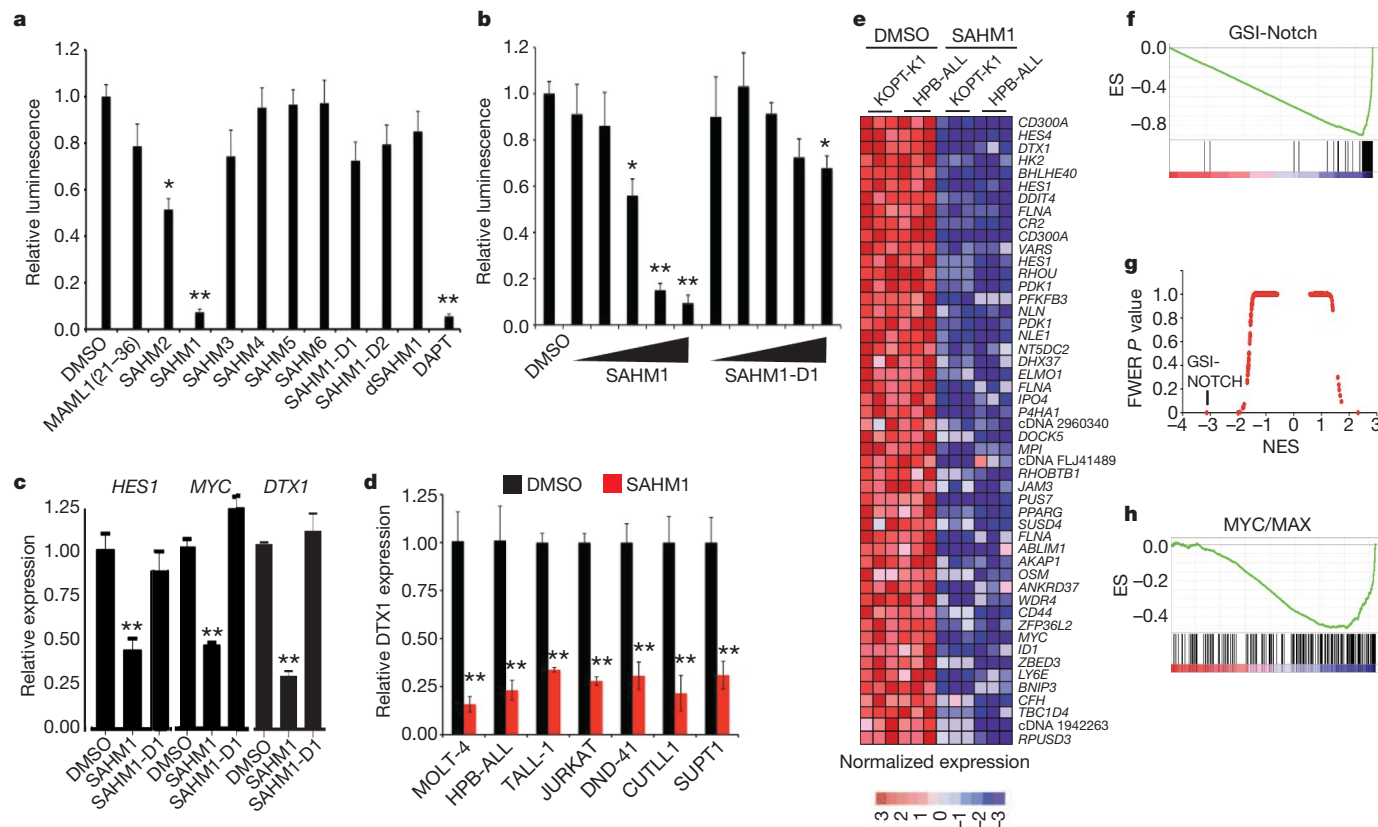
**Figure 3 | SAHM1 represses NOTCH1 target gene expression. a**, Inhibition of a NOTCH1-dependent luciferase reporter by SAHM peptides. Signal was normalized to *Renilla* luciferase control. **b**, Dose-dependent effects of SAHM1 and SAHM1-D1 in the dual-luciferase assay, using threefold dilutions (0.55–45 μM) of ligand compared to vehicle alone. **c**, qRT–PCR analysis of the *HES1*, *MYC* and *DTX1* mRNA levels in KOPT-K1 cells treated for 24 h with SAHM1 or SAHM1-D1 (20 μM) relative to dimethylsulphoxide (DMSO) control. **d**, qRT–PCR analysis of *DTX1* mRNA levels in a panel of human T-ALL cell lines. **e**, Heat map representation of the top 50 downregulated genes ($P < 0.001$), induced by SAHM1 in KOPT-K1 and HPB-ALL cells. **f**, Quantitative comparison of genes downregulated by GSI (GSI-NOTCH gene set) with the SAHM1 gene expression profile in KOPT-K1 and HPB-ALL cells by GSEA. **g**, Comparison of all transcription factor target gene sets in the Molecular Signatures Database to the GSI-NOTCH gene set for enrichment in the SAHM1 expression profile by GSEA. Data are plotted as the family-wise error rate (FWER) *P* value versus the NES. GSI-NOTCH is marked as the most enriched gene set. **h**, GSEA of the second most enriched gene set (MYC/MAX), applied to the SAHM1 expression profile. Unless noted otherwise data represent the mean ± s.d. ($n = 3$). $*P < 0.05$, $**P < 0.01$.

normalized ES (NES) $= -3.66$, $P < 0.0001$, Fig. 3f). Leading-edge analysis identified further enrichment for transcripts annotated as NOTCH1 targets in the scientific literature (Supplementary Fig. 4a). As a measure of biological specificity, enrichment analysis was conducted on all transcription factor target gene sets available in the Molecular Signatures Database (MSigDB). Taken together, the GSI-NOTCH gene set emerged as a statistical outlier as the most enriched in the SAHM1 profile (Fig. 3g and Supplementary Table 2). Interestingly, analysis of SAHM1-repressed genes also identified significant enrichment for gene sets regulated by transcriptional activators such as MYC/MAX and E2F, which have been previously identified as downstream targets of NOTCH activation (Fig. 3h and Supplementary Table 2)[40,41]. Taken together, these data establish that SAHM1 exerts a specific antagonistic effect on gene expression driven by NOTCH.

## NOTCH inhibition halts the proliferation of T-ALL cells

Previous work has shown that chemical (GSIs) or genetic (dnMAML1 expression) inhibition of NOTCH induces cell-cycle arrest, apoptosis and decreased proliferative capacity in a subset of T-ALL cell lines[29,42]. We assembled a panel of genetically annotated T-ALL cell lines, which have previously demonstrated sensitivity to GSI (CUTLL1, SUPT1, HPB-ALL, TALL-1, DND-41 and KOPT-K1; Supplementary Table 1). SAHM1 treatment in these cells produced a marked reduction in cell proliferation whereas SAHM1-D1 was inactive (Fig. 4a). SAHM1 treatment did not affect the proliferation of K562 cells, an erythroleukaemia

cell line without dependency on NOTCH1 for growth, or JURKAT and MOLT-4 cells, both of which are T-cell lines bearing activating NOTCH1 mutations but are insensitive to NOTCH inhibition owing to the loss of PTEN[15,41,43]. In sensitive T-ALL cell lines, SAHM1 exposure prompted activation of caspase 3 and 7, consistent with the induction of apoptosis (Fig. 4b).

These findings suggested that even brief exposure of T-ALL cells to SAHM1 might be sufficient to prevent the establishment of leukaemia *in vivo*. To recapitulate the human disease genotype and phenotype, we devised a murine model of T-ALL induced by the *NOTCH1* allele L1601PΔP, which bears dual NOTCH1 mutations originally identified in human T-ALL cells (Supplementary Fig. 5). Retroviral infection of the L1601PΔP allele into whole bone marrow followed by transplantation into lethally irradiated syngeneic recipient mice resulted in development of a T-ALL phenotype; primary cells demonstrated sensitivity to SAHM1 treatment *in vitro* (Supplementary Fig. 6a–c). To determine whether SAHM1 treatment could curb leukaemic engraftment, we pre-treated primary L1601PΔP cells with SAHM1 or vehicle before inoculation into secondary recipient mice (Supplementary Fig. 6d). At study end point, the SAHM1-treated cohort showed a statistically significant reduction in spleen weight (Fig. 4c) and the absolute number of donor-derived circulating lymphoblasts (Fig. 4d). Histopathological examination of bone marrow and spleen demonstrated a marked reduction in disease burden and leukaemic infiltration among animals administered SAHM1-treated cells (Fig. 4e).
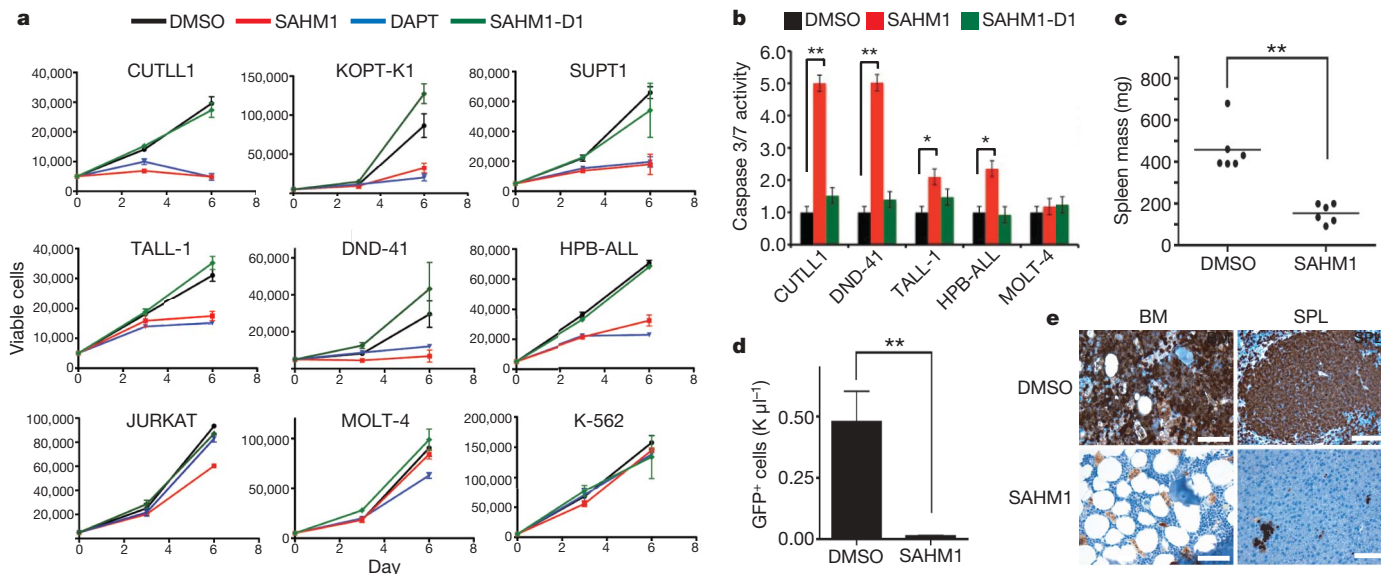
**Figure 4 | SAHM1 reduces T-ALL proliferation and leukaemic initiation potential. a,** Growth effects of SAHMs on a panel of human T-ALL cell lines of known mutational status. Cells were incubated with 15 μM DAPT (blue triangles), SAHM1 (red squares), SAHM1-D1 (green diamonds) or DMSO (black circles) and monitored for proliferation after 3 and 6 days of culture. Data points are mean ± s.d. (n = 3). **b,** Effects of SAHMs on apoptosis of T-ALL cells monitored using Capase-glo 3/7 (Promega) in cultures carried out as in **a.** Error bars, s.d. **c, d,** Ex-vivo treatment of L1601PΔP cells with SAHM1 (5 μM, 12 h) limits leukaemia initiation in secondary murine

recipients. Reduction of spleen weight in the SAHM1 cohort (n = 6) compared to vehicle (n = 6) at the first sign of disease toxicity (23 days) (P = 0.001) (**c**). Circulating GFP-positive cell count in the blood is reduced by ~100-fold in the SAHM1 cohort (n = 6) relative to vehicle (n = 6) (P = 0.0026) (**d**). Error bars, s.d. Statistical analyses performed with a two-tailed t-test. *P < 0.01, **P < 0.005. **e,** An immunohistochemical stain for GFP that imparts a brown colour shows extensive leukaemic infiltration of bone marrow (BM) and spleen (SPL) in representative mice receiving vehicle-treated transplants relative to SAHM1. SPL, spleen. Scale bars,

## SAHM1 inhibits leukaemic progression and NOTCH1 signalling

To evaluate whether SAHM1 treatment could attenuate the growth of established T-ALL in vivo, we developed a bioluminescent murine model of T-ALL. Either MIG-L1601PΔP or the control MIG retro-virus was used to infect bone marrow isolated from C57BL/6 mice in which firefly luciferase was constitutively expressed from the ubiquitin C promoter[44]. Infected cells were then transplanted into isogenic C57BL/6-Tyr$^{C/C}$ albino recipients to facilitate non-invasive bio-luminescence imaging[45]. Animals transplanted with Luc-L1601PΔP cells developed progressive T-ALL (Fig. 5a) that was quantified by serial imaging (Fig. 5b), whereas no measurable disease developed in mice receiving control-transduced cells (Fig. 5a, b).

Balanced cohorts of secondary recipient mice with established leukaemia were treated with vehicle, daily SAHM1 (35 mg kg$^{-1}$), or twice-daily SAHM1 (30 mg kg$^{-1}$) by intraperitoneal injection. Vehicle-treated mice showed progressive disease, with eight out of nine mice showing increased bioluminescence over 5 days of treatment (Fig. 5c). Mice receiving daily SAHM1 treatments showed a slightly lower mean change in bioluminescence and included fewer animals with progressive disease (four out of six). All mice treated with twice-daily SAHM1, however, showed a significant dose-dependent regression of tumour as evidenced by a decrease in bio-luminescence.

To confirm that the anti-leukaemic effect observed with SAHM1 treatment was associated with attenuated NOTCH1 signalling, mononuclear cells from vehicle- and SAHM1-treated (twice-daily) animals were collected for NOTCH1 target gene transcriptional analysis. A significant decrease in messenger RNA levels was evident for *Heyl, Hes1, Myc, Dtx1* and *Nrarp* in mice treated with SAHM1 compared to vehicle (Fig. 5d). Gene expression profiling also revealed significant pharmacodynamic repression of the NOTCH transcriptional program *in vivo*. A gene set of murine NOTCH1 targets was curated from a published report in which NOTCH1-dependent murine T6E cells were profiled after treatment with GSI or ectopic expression of dnMAML1 (ref. 40). Significant enrichment for the murine NOTCH1 gene set was observed in isolated cells from

SAHM1-treated animals (Fig. 5e and Supplementary Fig. 7). Together, these data provide a direct link between inhibition of the NOTCH pathway and the anti-leukaemic activity of SAHM1 *in vivo*.

## Discussion

Chemical intractability has limited the discovery of synthetic entities targeting transcription factors. Here we set out to assess whether a new type of targeting molecule, a stapled peptide derived from MAML1, could target a seemingly intractable transcription factor, human NOTCH1. We demonstrated direct binding to the pre-assembled form of the NOTCH1–CSL complex and competitive inhibition of MAML1 co-activator binding. Analysis of direct NOTCH1 target gene levels and the global expression profile induced by SAHM1 confirmed specific repression of the NOTCH signalling program in human and murine T-ALL cells. Given the complexities of transcriptional responses and the state of their characterization, it is impossible to rule out conclusively any off-target activity by SAHM1. In the context of human and murine T-ALL however, GSEA provides a striking correlation between the expression effects of SAHM1 and a GSI. Thus, it is clear that the NOTCH pathway is the major target of SAHM1. Direct transcriptional repression was subsequently found to enact NOTCH-specific anti-proliferative effects in annotated T-ALL cell lines. Finally, SAHM1 treatment curbed leukaemic progression and inhibited NOTCH1 signalling in a relevant murine model of T-ALL. Previous studies in mice and preliminary data in humans have shown that treatment with GSI alone is toxic owing to on-target effects on intestinal crypts, suggesting that different targeting strategies or drug combinations are needed; indeed recent reports are showing promise[46]. Our studies so far have not observed gastrointestinal toxicity at necropsy in treated animals (Supplementary Fig. 8a, b). Although expanded evaluation of SAHM1 as a possible therapeutic agent is needed, these early results indicate a potential therapeutic window.

As a direct transcriptional antagonist, SAHM1 should prove broadly useful in further determining the role of NOTCH in both normal tissues and disease processes, and presents a starting-point
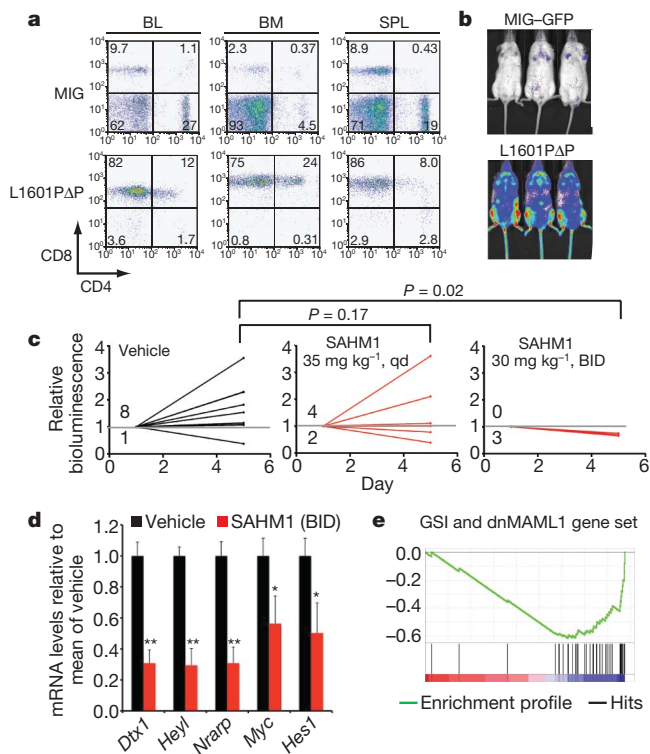
**Figure 5 | SAHM1 treatment inhibits NOTCH signalling and leukaemic progression *in vivo*. a,** Flow cytometric analysis of cells isolated from C57BL/6-Tyr$^{C/C}$ mice reconstituted with luciferase-expressing haematopoietic stem cells transduced with an L1601PΔP *NOTCH1* allele or empty vector (MIG). BL, blood. **b,** Bioluminescence imaging of primary recipients of Luc-L1601PΔP cells or control cells 2 months post-reconstitution. **c,** Bioluminescence quantification of tumour burden in mice with established disease treated with vehicle, 35 mg kg$^{-1}$ SAHM1 per day (QD; $P = 0.17$), or 30 mg kg$^{-1}$ SAHM1 twice daily (BID; $P = 0.02$). A Fisher's Exact test was used to compare disease progression between the cohorts. **d,** qRT–PCR analysis reveals repression of *Hes1* ($P = 0.0187$), *Myc* ($P = 0.023$), *Nrarp* ($P = 0.001$), *Heyl* ($P = 0.0006$) and *Dtx1* ($P = 0.0006$) mRNA levels in blood collected at day five from vehicle ($n = 3$) and SAHM1 (30 mg kg$^{-1}$ BID, $n = 3$)-treated mice. **e,** Enrichment of GSI and dnMAML1 downregulated transcripts in the gene expression profile of leukocytes isolated from SAHM1-treated mice. Data in **d** represent the mean ± s.d. of triplicate measurements. Unless otherwise noted statistical analyses were performed with a two-tailed *t*-test. *$P < 0.05$, **$P < 0.005$.

for the development of a targeted therapeutic agent to treat NOTCH-driven cancers and non-malignant conditions. Furthermore, we expect the approach described here will prove broadly applicable to several other transcription factor complexes previously considered beyond the reach of ligand discovery.

## METHODS SUMMARY

SAHM peptides were synthesized manually using standard 9-fluorenylmethoxy-carbonyl (Fmoc)-peptide chemistry on MBHA rink amide resin, and cross-linked using Grubbs-I catalyst (benzylidene-bis(tricyclohexylphosphine)dichlororuthenium). Recombinant human CSL (amino acids 9–435), GST–RAMANK and RAMANK (amino acids 1761–2127 of NOTCH1) and dnMAML1 (amino acids 13–74) were expressed and purified as previously reported[47]. GST pull-down and fluorescence polarization experiments were performed in 20 mM Tris, pH 8.4, containing 150 mM NaCl, 1 mM EDTA and 1 mM dithiothreitol (DTT). SPR experiments were performed in the identical buffer with 187 mM NaCl and 0.01% P-20. Fluorescence polarization was measured at $\lambda_{ex} = 485$ nm and $\lambda_{em} = 525$ nm on a Perkin Elmer Spectramax-M5 multi-label plate reader. SPR experiments were performed on a Biacore S51 SPR instrument using anti-GST and streptavidin functionalized CM5 chips. Luciferase and β-lactamase reporter gene assays were performed as described[35,36]. For quantitative RT–PCR (qRT–PCR), RNA was extracted from T-ALL cells, reverse transcribed to complementary DNA, and amplified using standard primers and probes specific for human and murine

transcripts. Expression values were calculated by the $2^{-\Delta\Delta C_t}$ method relative to the β-actin gene. RNA from triplicate SAHM- and vehicle-treated KOPT-K1 and HPB-ALL samples was used for analysis on Affymetrix U133 Plus 2.0 expression arrays, processed using GenePattern software. Gene set enrichment analysis was performed using GSEA software (http://www.broad.mit.edu/GSEA). Cell viability assays were performed in white, 96-well plates (Corning) in RPMI media containing 10% FCS. Fresh media and compounds were added every 3 days as needed. The L1601PΔP NOTCH1-dependent murine T-ALL cells were generated by reconstitution of lethally irradiated syngeneic mice with bone marrow cells transduced with the L1601PΔP NOTCH1 allele as described[48]. Luc-L1601PΔP animals were regularly monitored for bioluminescence, and euthanized at 2–3 months after transplant to obtain splenic T-ALL cells for secondary transplants and *in vivo* studies. Secondary recipients were monitored for leukaemia by bioluminescence and subsequently separated into treatment cohorts with matched disease burden.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Darnell, J. E. Jr. Transcription factors as targets for cancer therapy. *Nature Rev. Cancer* 2, 740–749 (2002).
2. Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* 284, 770–776 (1999).
3. Bray, S. J. Notch signalling: a simple pathway becomes complex. *Nature Rev. Mol. Cell Biol.* 7, 678–689 (2006).
4. Struhl, G. & Greenwald, I. Presenilin is required for activity and nuclear access of Notch in *Drosophila*. *Nature* 398, 522–525 (1999).
5. Ye, Y., Lukinova, N. & Fortini, M. E. Neurogenic phenotypes and altered Notch processing in *Drosophila* Presenilin mutants. *Nature* 398, 525–529 (1999).
6. De Strooper, B. *et al.* A presenilin-1-dependent γ-secretase-like protease mediates release of Notch intracellular domain. *Nature* 398, 518–522 (1999).
7. Fortini, M. E. & Artavanis-Tsakonas, S. The suppressor of hairless protein participates in notch receptor signaling. *Cell* 79, 273–282 (1994).
8. Petcherski, A. G. & Kimble, J. Mastermind is a putative activator for Notch. *Curr. Biol.* 10, R471–R473 (2000).
9. Wu, L. *et al.* MAML1, a human homologue of *Drosophila* Mastermind, is a transcriptional co-activator for NOTCH receptors. *Nature Genet.* 26, 484–489 (2000).
10. Li, L. *et al.* Alagille syndrome is caused by mutations in human *Jagged1*, which encodes a ligand for Notch1. *Nature Genet.* 16, 243–251 (1997).
11. Joutel, A. *et al. Notch3* mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* 383, 707–710 (1996).
12. Garg, V. *et al.* Mutations in *NOTCH1* cause aortic valve disease. *Nature* 437, 270–274 (2005).
13. Ellisen, L. W. *et al.* TAN-1, the human homolog of the *Drosophila notch* gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* 66, 649–661 (1991).
14. Weng, A. P. *et al.* Activating *mutations* of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 306, 269–271 (2004).
15. O'Neil, J. *et al.* FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to γ-secretase inhibitors. *J. Exp. Med.* 204, 1813–1824 (2007).
16. Pece, S. *et al.* Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis. *J. Cell Biol.* 167, 215–221 (2004).
17. Fre, S. *et al.* Notch signals control the fate of immature progenitor cells in the intestine. *Nature* 435, 964–968 (2005).
18. Uyttendaele, H. *et al. Notch4/int-3*, a mammary proto-oncogene, is an endothelial cell-specific mammalian *Notch* gene. *Development* 122, 2251–2259 (1996).
19. Nicolas, M. *et al.* Notch1 functions as a tumor suppressor in mouse skin. *Nature Genet.* 33, 416–421 (2003).
20. Konishi, J. *et al.* γ-secretase inhibitor prevents Notch3 activation and reduces proliferation in human lung cancers. *Cancer Res.* 67, 8051–8057 (2007).
21. Nefedova, Y., Cheng, P., Alsina, M., Dalton, W. S. & Gabrilovich, D. I. Involvement of Notch-1 signaling in bone marrow stroma-mediated de novo drug resistance of myeloma and other malignant lymphoid cell lines. *Blood* 103, 3503–3510 (2004).
22. Park, J. T. *et al. Notch3* gene amplification in ovarian cancer. *Cancer Res.* 66, 6312–6318 (2006).
23. Miyamoto, Y. *et al.* Notch mediates TGFα-induced changes in epithelial differentiation during pancreatic tumorigenesis. *Cancer Cell* 3, 565–576 (2003).
24. Seiffert, D. *et al.* Presenilin-1 and -2 are molecular targets for γ-secretase inhibitors. *J. Biol. Chem.* 275, 34086–34091 (2000).
25. Dovey, H. F. *et al.* Functional γ-secretase inhibitors reduce β-amyloid peptide levels in brain. *J. Neurochem.* 76, 173–181 (2001).
26. Lleó, A. Activity of γ-secretase on substrates other than APP. *Curr. Top. Med. Chem.* 8, 9–16 (2008).
27. Riccio, O. *et al.* Loss of intestinal crypt progenitor cells owing to inactivation of both Notch1 and Notch2 is accompanied by derepression of CDK inhibitors p27$^{Kip1}$ and p57$^{Kip2}$. *EMBO Rep.* 9, 377–383 (2008).
28. Maillard, I. *et al.* Mastermind critically regulates Notch-mediated lymphoid cell fate decisions. *Blood* 104, 1696–1702 (2004).

29. Weng, A. P. et al. Growth suppression of pre-T acute lymphoblastic leukemia cells by inhibition of notch signaling. Mol. Cell. Biol. 23, 655–664 (2003).

30. Nam, Y., Sliz, P., Song, L., Aster, J. C. & Blacklow, S. C. Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes. Cell 124, 973–983 (2006).

31. Wilson, J. J. & Kovall, R. A. Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA. Cell 124, 985–996 (2006).

32. Schafmeister, C. E., Po, J. & Verdine, G. L. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. J. Am. Chem. Soc. 122, 5891–5892 (2000).

33. Walensky, L. D. et al. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. Science 305, 1466–1470 (2004).

34. Del Bianco, C., Aster, J. C. & Blacklow, S. C. Mutational and energetic studies of Notch 1 transcription complexes. J. Mol. Biol. 376, 131–140 (2008).

35. Aster, J. C. et al. Essential roles for ankyrin repeat and transactivation domains in induction of T-cell leukemia by notch1. Mol. Cell. Biol. 20, 7505–7515 (2000).

36. Hallis, T. M. et al. An improved β-lactamase reporter assay: multiplexing with a cytotoxicity readout for enhanced accuracy of hit identification. J. Biomol. Screen. 12, 635–644 (2007).

37. Lamb, J. et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313, 1929–1935 (2006).

38. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl Acad. Sci. USA 102, 15545–15550 (2005).

39. Palomero, T. et al. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc. Natl Acad. Sci. USA 103, 18261–18266 (2006).

40. Weng, A. P. et al. c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev. 20, 2096–2109 (2006).

41. Rao, S. S. et al. Inhibition of NOTCH signaling by gamma secretase inhibitor engages the RB pathway and elicits cell cycle exit in T-cell acute lymphoblastic leukemia cells. Cancer Res. 69, 3060–3068 (2009).

42. Lewis, H. D. et al. Apoptosis in T cell acute lymphoblastic leukemia cells after cell cycle arrest induced by pharmacological inhibition of notch signaling. Chem. Biol. 14, 209–219 (2007).

43. Palomero, T. et al. Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. Nature Med. 13, 1203–1210 (2007).

44. Becker, C. M. et al. A novel noninvasive model of endometriosis for monitoring the efficacy of antiangiogenic therapy. Am. J. Pathol. 168, 2074–2084 (2006).

45. Stubbs, M. C. et al. MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia: development of a model for rapid therapeutic assessment. Leukemia 22, 66–77 (2008).

46. Real, P. J. et al. γ-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia. Nature Med. 15, 50–58 (2009).

47. Nam, Y., Weng, A. P., Aster, J. C. & Blacklow, S. C. Structural requirements for assembly of the CSL·intracellular Notch1·Mastermind-like 1 transcriptional activation complex. J. Biol. Chem. 278, 21232–21239 (2003).

48. Chiang, M. Y. et al. Leukemia-associated NOTCH1 alleles are weak tumor initiators but accelerate K-ras-initiated leukemia. J. Clin. Invest. 118, 3181–3194 (2008).

# METHODS

**Chemical reagents.** Solid phase rink amide MBHA resin (100–200 mesh), 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), natural Fmoc-protected L-amino acids and N-biotinyl-NH-PEG$_2$-COOH were all purchased from Novabiochem (EMD Chemicals). The γ-secretase inhibitor N-[N-(3,5-difluorophenacetyl-L-alanyl)]-S-phenyl-glycine t-butyl ester (DAPT) was purchased as a 25 mM solution in DMSO from Calbiochem (EMD Chemicals). All other reagents and solvents for chemical synthesis were purchased from Sigma Aldrich.

**Solid-phase peptide synthesis.** All peptides were synthesized manually using standard Fmoc-peptide chemistry on MBHA rink amide resin. Chemical cross-linking of non-natural amino acids containing pentenyl side-chains was performed using 0.1 equivalents of Grubbs-I catalyst (benzylidene-bis(tricyclo-hexylphosphine)dichlororuthenium) dissolved in dichloroethane under Argon for 2.5 h. After olefin-metathesis, peptides were subjected to one final round of deprotection and capping by standard methods with acetic anhydride, FITC or N-biotinyl-NH-PEG$_2$-COOH creating SAHM, FITC–SAHM and bioSAHM peptides, respectively. After completion, crude peptide was then dissolved in 1:1 acetonitrile:water, and purified by reverse-phase high-pressure liquid chromatography (HPLC) using a C18 column (Agilent). Compound identification and purity was assessed using coupled liquid chromatography mass spectrometry (LCMS) (Agilent).

Synthesis of non-natural amino acids was performed as reported[49].

**Determination of α-helicity.** A circular dichroism spectrometer (Jasco J-710) was used to determine the α-helical content of all SAHM peptides. SAHMs were dissolved into deionized water at pH 7.4 to a final concentration of 25–100 μM in a quartz cuvette with a path length of 1 cm. Absorbance values were taken at 1 nm intervals between 190 and 255 nm. The percentage helicity was calculated from the absorbance at 222 nm using helical models as previously reported in ref. 50.

**Protein expression and purification.** Human CSL (residues 9–435) was expressed with a carboxy-terminal hexahistidine tag in *Escherichia coli* strain BL21(DE3) pLysS (Stratagene) using a pET28a vector (Novagen). Transformed bacteria was grown at 18 °C and induced with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) at an $A_{600} = 0.8$. Cells were pelleted 14 h after induction and lysed in buffer 1 (50 mM Tris, pH 8.4, 500 mM NaCl, 2 mM DTT, 1 mM EDTA and 1 mM phenylmethyl-sulphonyl fluoride (PMSF)) by sonication. The lysate was centrifuged to clear insoluble matter before loading onto Ni-NTA resin (Qiagen). Column-bound protein was eluted using 250 mM imidazole in buffer 1 and concentrated by centrifugation using a 30-kDa exclusion filter. The concentrated eluent was dialysed into buffer 2 (20 mM Tris, pH 8.4, 1 mM DTT, 1 mM EDTA and 150 mM NaCl) before further purification by gel filtration on a Superdex-75 Column (Amersham Pharmacia Biotech). Human RAMANK (amino acids 1761–2127) of ICN1 was expressed and purified as reported[47]. The corresponding GST-fusion protein was bound to glutathione-Sepharose beads, which were washed with buffer 1 without protease inhibitors. The protein was then eluted from the beads with a solution of 20 mM glutathione in buffer 1 to obtain GST–RAMANK, or cleaved directly from the resin using TEV protease to obtain RAMANK. Both proteins were further purified by gel filtration.

**GST pull-down assays.** Glutathione-functionalized agarose beads (GST-Bind, Biorad) were incubated with saturating concentrations of purified GST–RAMANK in buffer 2 for 30 min at 4 °C. Beads were washed and 40 μl of the 50% bead slurry was incubated with indicated mixtures of purified CSL (0 or 0.5 μM), dnMAML1 (0, 0.5, 2.5 or 5 μM) and SAHM1 (0 or 10 μM) to a final volume of 350 μl, and incubated for 1 h at 4 °C. Beads were washed four times with buffer 2, taken up in SDS–PAGE loading buffer, boiled, and SDS–PAGE was performed to visualize bound proteins by Coomassie staining.

**Fluorescence polarization assays.** SAHM-binding assays were performed by incubating FITC–SAHM peptides (15 nM) with twofold dilutions of equimolar RAMANK and CSL protein in buffer 2. Dilutions and incubations were made in 384-well, black flat-bottom plates (Corning) to a total volume of 40 μl, and equilibrated for 30 min. Polarization was measured on a Spectramax-M5 multi-label plate reader with $\lambda_{ex} = 485$ nm and $\lambda_{em} = 525$ nm. Polarization was calculated according to the standard equation: $P = (V - H)/(V + H)$, in which $P$ denotes polarization, $V$ denotes vertical emission intensity, and $H$ represents horizontal emission intensity. $K_d$ values were determined by fitting data to a variable-slope sigmoidal binding curve using Prism 4 graphing software. Competitive fluorescence polarization assays were performed by incubating fourfold dilutions of dnMAML1 or DMSO alone with FITC–SAHM1 (15 nM) and a pre-equilibrated mixture of RAMANK–CSL complex (600 nM). Polarization measurements and calculations were performed as above for binding assays.

**Surface plasmon resonance.** A Biacore S51 SPR-Instrument (Biacore-GE) was used for all SPR-binding measurements. Binding between the RAMANK domains of ICN1 and CSL was performed by first immobilizing IgG anti-GST (Pierce) on a CM5 Biacore chip by amine coupling with EDC (1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride) activation and NHS (N-hydroxysuccinimide) coupling. All proteins were exchanged into binding buffer (187 mM NaCl, 20 mM Tris, pH 8.4, 1 mM DTT and 0.05% P-20) before use. GST–RAMANK (50 μg ml$^{-1}$) was injected at a rate of 10 μl min$^{-1}$ for 10 min, immobilizing approximately 400 response units. Escalating concentrations of CSL (twofold dilutions between 4.0 and 0.125 μM, including two blanks) were subsequently passed over the GST–RAMANK functionalized chip as well as an anti-GST reference surface. After each CSL dilution the sensor surface was regenerated with pH 2.2 Tris-glycine solution and again functionalized with a comparable level of GST–RAMANK to avoid carry-over between experiments. Binding data was reference normalized and processed using ClampXP software (http://www.cores.utah.edu/interaction/clamp.html). A two-site binding model was applied to the processed data set to determine kinetic parameters of the RAMANK–CSL interaction. SAHM binding to the RAMANK–CSL complex was performed using SAHM1 and SAHM1-D1 each capped at the N terminus with N-biotinyl-NH-(PEG)$_2$-COOH (Novabiochem). Peptides were dissolved in binding buffer and injected at 10 μl min$^{-1}$ for 10 min into a flow cell containing a Streptavidin-CM5 Biacore chip. Equimolar dilutions of RAMANK and CSL (twofold dilutions from 1 to 0.03125 μM, including two blanks) were mixed in binding buffer and injected onto the peptide-functionalized surface to measure their association and dissociation to SAHM1 and SAHM1-D1. As with the RAMANK–CSL interaction, kinetic and thermodynamic data was generated using a two-site binding model in ClampXP software.

**Western blots and co-immunoprecipitation assays.** For ligand pull-down assays, biotinylated SAHMs were immobilized on Streptavidin-agarose beads (30 μl slurry) in 250 μl lysis buffer for 4 h at 4 °C. Separately, $2 \times 10^6$ KOPT-K1 cells were lysed in 20 mM Tris, pH 8, 150 mM NaCl, 1% Triton X-100, 0.1 mM NaF, 1 mM PMSF, 1 mM sodium orthovanadate and 2.5 μg ml$^{-1}$ leuptin supplemented with complete-mini protease inhibitor (Roche) by sonication, followed by removal of insoluble debris by centrifugation at 4 °C. Clarified lysates were added to the SAHM-immobilized beads and incubated overnight. Beads were washed repeatedly with lysis buffer before the elution of bound proteins by boiling in gel-loading buffer for 10 min. Eluted proteins were processed for western blotting by standard methods using polyclonal antibodies specific for ICN1 (Val1744, Cell Signaling) or CSL (CBF-1, Santa Cruz Bio.). MAML1 immunoprecipitations were performed and analysed in a similar manner. Antibodies directed at MAML1 (Cell Signaling Technologies) were immobilized on Protein-A agarose beads for 4 h and washed with lysis buffer. Separately, cellular lysates were incubated with soluble FITC–SAHM in a total volume of 250 μl for 4 h at 4 °C. Lysates were then added to antibody-bound beads, incubated overnight and processed for western blot analysis as described earlier.

**Fluorescence microscopy.** High content imaging measurements were made using an IXMicro epifluorescence microscope (Molecular Devices Corporation) and quantified using MetaXpress image analysis software. To determine intracellular penetration, $2 \times 10^3$ cells were seeded in black, clear-bottom 384-well imaging plates (Costar) and incubated in 40 μl RPMI (KOPT-K1 cells) or DMEM (U2OS cells) media supplemented with 10% FBS and the compounds of interest. SAHM peptides (5 μM) were incubated with U2OS cells for 12 h at 37 °C. Cells were then washed three times with PBS, fixed with 3.7% formaldehyde in PBS, and counterstained with Hoescht dye (1:5,000). Epifluorescent images were taken in triplicate, imported into MetaXpress software, and scored as positive or negative on the basis of minimum size, maximum size, fluorescence intensity above local background and overlap of Hoescht and FITC signals. Rhodamine-labelled SAHM1 was visualized and scored in the same manner. The temperature dependence of SAHM1 penetration was determined in U2OS cells incubated with 5 μM FITC–SAHM1 for 4 h at either 37 °C or 4 °C as earlier.

**Luciferase reporter gene assays.** CSL-dependent luciferase reporter assays were performed as reported[35].

**β-lactamase reporter gene assays.** HeLa cells containing an integrated CSL-responsive β-lactamase transgene and a doxycycline-inducible ICN1 transgene were generously provided by Invitrogen. Cells were maintained under selection in DMEM containing 10% FBS, 5 μg ml$^{-1}$ blasticidin, 125 μg ml$^{-1}$ hyrgromycin and 75 μg ml$^{-1}$ zeocin. Cells were seeded in 384-well imaging plates and ICN1 was induced with 150 ng ml$^{-1}$ doxycycline for 24 h. After induction, stock solutions of SAHM peptides, potassium clavulanate or DMSO were added and cells were incubated for a further 24 h. The β-lactam fluorescence resonance energy transfer (FRET) substrate loading solution was prepared separately according to the manufacturer's specifications (ToxBlaZer substrate, Invitrogen Corp.). FRET loading solution (5 μl) was added to each induced well, and also to non-induced and cell-free wells to determine basal FRET ratio levels and fluorescence background, respectively. After 2 h, fluorescence measurements were made using

a Spectramax-M5 multi-label plate reader (Molecular Devices). The FRET probe was excited at 409 nm, and the emission values at 460 nm (cleaved probe) and 530 nm (intact probe) were measured. The background fluorescence of cell-free wells at both wavelengths was subtracted before calculating the 460 nm/530 nm emission ratio.

**Quantitative RT–PCR.** One-million human T-ALL cells were seeded in 24-well plates in RPMI containing 10% FCS. SAHM peptides (20 μM) were added and cells were then incubated for 24 h at 37 °C. RNA samples were prepared using an RNEasy Mini kit (Qiagen). Total RNA was reverse transcribed to cDNA using SuperScript II Reverse Transcriptase Kit according to the manufacturer's protocol (Invitrogen). The resulting cDNA was used as the substrate to measure relative expression levels by qRT–PCR with TaqMan (Applied Biosystems) probes and primers specific for human *HES1*, *MYC*, *DTX1* and *ACTB* for normalization: probe IDs Hs00172878_m1, Hs00153408_m1, Hs01092201_m1 and Hs99999903_m1, respectively. In other experiment involving murine T-ALL cells, the following TaqMan probes and primers specific for murine *Hes1*, *Myc*, *Dtx1*, *Nrarp*, *Heyl* and *Actb* were used: Mm01342805_m1, Mm00487807_m1, Mm00492297_m1, Mm00482529_s1, Mm00516555_m1 and Mm01205647_m1, respectively. Triplicate reactions were prepared in 384-well optical PCR plates (Applied Biosystems) and qPCR was performed using a ABI Prism 7900 (Applied Biosystems) instrument with the following run specifications: FAM probe detection, AmpErase activation for 2 min at 50 °C, DNA polymerase activation for 10 min at 95 °C, followed by 40 cycles of duplex melting for 15 s at 95 °C and annealing for 1 min at 60 °C. Threshold-cycle ($C_t$) values were automatically calculated for each replicate and used to determine the relative expression of the gene of interest relative to reference genes for both treated and untreated samples by the $2^{-\Delta\Delta Ct}$ method.

**Gene expression profiling.** One-to-two micrograms of RNA per array was prepared from cultures of KOPT-K1 and HPB-ALL cells treated for 24 h with 20 μM SAHM1 or equivalent DMSO as described earlier for qRT–PCR in triplicate. Expression profiles were then generated by hybridizing processed RNA with Human Genome U133 Plus 2.0 arrays (Affymetrix). Murine T-ALL gene expression profiles were generated with RNA isolated from vehicle and SAHM1-treated mice using Affymetrix Mouse 430 2.0 arrays. cDNA processing, chip preparation, hybridization and chip scanning were performed by the Broad Institute Microarray Core Facility. Raw .cel files were processed and normalized by the robust multiarray averaging (RMA) method using GenePattern software (http://www.broad.mit.edu/cancer/software/genepattern/). For the KOPT-K1 and HPB-ALL samples, RMA processed data was subsequently normalized to the mean of DMSO-treated samples on a per gene basis to control for interarray differences. The GSI-NOTCH gene set was compiled from the published GSI study by extracting the top gene identifiers consistently downregulated across seven T-ALL cell lines by signal-to-noise ratio. This list of probes was applied to the SAHM1 expression profile by GSEA software (http://www.broad.mit.edu/GSEA/). For the global analysis of all transcription factor target gene sets (C3 TFT gene sets) we downloaded a .gmt file containing all gene sets and added in the curated GSI-NOTCH gene set. This resulting .gmt database was queried against the SAHM1 expression profile with the following parameters: probe set collapse = true; phenotype = SAHM1 versus DMSO; permutation: gene set, permutations = 1,000. Gene set size: $15 < n < 500$. The MYC/MAX gene set in Fig. 3h is the second most enriched gene set titled 'V$MYCMAX_01' in the MSigDB.

**Cell proliferation and apoptosis assays.** Five-thousand cells were seeded in white, 96-well plates (Corning) in a total volume of 125 μl RPMI media containing 1% penicillin/streptomycin and 10% FBS. Media also contained 15 μM DAPT, SAHM1, SAHM1-D1 or equivalent concentrations of DMSO (0.3%). Cells were incubated at 37 °C for 72 h; at this point, the media was changed and fresh compounds added. The viable cell number was determined at 72 h and 144 h using the Cell Titer-Glo viability assay (Promega). All measurements were made in triplicate, and the growth curves shown are representative of several independent experiments. Apoptosis was determined at 144 h using the Caspase-Glo 3/7 assay (Promega), which reports on the activation of caspases 3 and 7 using a pro-luminescent substrate.

**Viral infection and bone marrow transplantation.** The MSCV-L1601P-ΔPEST-GFP (L1601PΔP) construct and the preparation of viral supernatants have been described[48]. Transduction efficiency in primary cells was confirmed by flow cytometric analysis of GFP expression. For bone marrow transplantation, 8–10-week-old wild-type Balb/c donor mice or $Tyr^{-/-}$, C57BL/6, Ubc-Luc mice were injected with 5-FU (Sigma) 6 days before bone marrow collection from femurs and tibiae. After an overnight incubation in RPMI-1640 supplemented with 10% FBS, 10 ng ml$^{-1}$ mouse IL-3, 20 ng ml$^{-1}$ mouse IL-6 and 10 ng ml$^{-1}$ mouse SCF, cells were spin-infected with viral supernatants twice over 2 days. Transduced cells ($1 \times 10^6$) were injected into the tail-vein of lethally irradiated syngeneic recipient mice. Animals were euthanized 2–3 months after transplant to obtain splenic T-ALL cells for secondary transplants and *in vivo* studies.

**Leukaemia initiation studies.** Primary murine T-ALL cells expressing L1601PΔP were re-suspended in 10% FBS-supplemented RPMI at a density of $1 \times 10^6$ cells per ml in 6-well plates. SAHM peptides or DMSO vehicle were added to a final concentration of 5 μM in 6 ml of media, and incubated with the cells for 12 h at 37 °C. Cells were then pelleted, resuspended in 3 ml of HBSS, and injected by tail vein ($1 \times 10^6$ cells per mouse) into two cohorts of syngeneic Balb/c mice ($n = 6$) that had been sub-lethally irradiated 12 h previously. The primary end point of this *in vivo* study was spleen weight; secondary end points include peripheral blood counts and GFP$^+$ blood counts just before necropsy. Approval for animal use in this study was granted by the Children's Hospital Boston Animal Care and Use Committee.

***In vivo* efficacy experiments.** Primary and secondary recipients were generated as above using Ubc-Luc C57BL/6-Tyr$^{C/C}$ mice. Secondary recipients were monitored for leukaemic engraftment and progression using bioluminescent imaging as previously described[45]. Approximately 2–3 weeks after reconstitution, mice showing established disease were divided into cohorts for treatment with vehicle alone (5% DMSO in HBS), 35 mg kg$^{-1}$ SAHM1 once per day, or 30 mg kg$^{-1}$ SAHM1 twice per day, given intraperitoneally. At the end of 5 days of treatment, mice were again imaged to determine leukaemic progression, and euthanized to examine the pharmacodynamic effects of the various treatments on NOTCH1 target-gene expression. Blood was collected by terminal cardiac bleed from mice treated with vehicle ($n = 3$) or SAHM1 (30 mg kg$^{-1}$ twice daily, $n = 3$) for qPCR and microarray gene expression analysis.

**Histopathological analysis and immunohistochemistry.** Paraffin-embedded tissue sections were prepared at the Dana Farber/Harvard Cancer Center Specialized Histopathology Services Core or Rodent Histopathology Core. In brief, tissues were fixed for at least 72 h in 10% neutral-buffered formalin (Sigma), dehydrated in alcohol, cleared in xylene, and infiltrated with paraffin on an automated processor (Leica). Tissue sections (4-μm thick) were placed on charged slides, deparaffinized in xylene, rehydrated through graded alcohol solutions, and stained with haematoxylin and eosin (H&E) or with anti-CD3 (Cell Marque), anti-B220 (BD Pharmingen), or anti-Ki-67 (Dako) antibodies. Images were obtained using an Olympus BX40 microscope (Olympus Imaging America) and a Micropublisher 3.3 RTV colour camera (QImaging).

**Flow cytometry.** All antibodies were obtained from BD Pharmingen and staining procedures were performed in PBS, 2% FBS. Freshly collected splenocytes or bone marrow cells ($0.5 \times 10^6$ per condition) were blocked with anti-CD16/32 antibody for 10 min and then incubated with 2 μl of each specific antibody for 30 min on ice in the dark. After two washes with PBS, 2% FBS, cells were resuspended in 200 μl PBS and analysed by flow cytometry. All analyses were gated on GFP$^+$ cells.

49. Williams, R. M. & Im, M. N. Asymmetric synthesis of monosubstituted and α, α-disubstituted α-amino acids via diastereoselective glycine enolate alkylations. *J. Am. Chem. Soc.* **113**, 9276–9286 (1991).

50. Chen, Y. H., Yang, J. T. & Chau, K. H. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry* **13**, 3350–3359 (1974).

# Enhanced lithium depletion in Sun-like stars with orbiting planets

Garik Israelian[1,2], Elisa Delgado Mena[1,2], Nuno C. Santos[3,4], Sergio G. Sousa[1,3], Michel Mayor[4], Stephane Udry[4], Carolina Domínguez Cerdeña[1,2], Rafael Rebolo[1,2,5] & Sofia Randich[6]

**The surface abundance of lithium on the Sun is 140 times less than the protosolar value[1], yet the temperature at the base of the surface convective zone is not hot enough to burn—and hence deplete—Li (refs 2, 3). A large range of Li abundances is observed[4,5] in solar-type stars of the same age, mass and metallicity as the Sun, but such a range is theoretically difficult to understand[3,6,7]. An earlier suggestion[8–10] that Li is more depleted in stars with planets was weakened by the lack of a proper comparison sample of stars without detected planets. Here we report Li abundances for an unbiased sample of solar-analogue stars with and without detected planets. We find that the planet-bearing stars have less than one per cent of the primordial Li abundance, while about 50 per cent of the solar analogues without detected planets have on average ten times more Li. The presence of planets may increase the amount of mixing and deepen the convective zone to such an extent that the Li can be burned.**

We obtained Li abundances from high-resolution, high signal-to-noise spectra for a sample of 451 stars in the HARPS high-precision (better than $1 \, \text{m s}^{-1}$) radial velocity exoplanet survey[11] spanning the effective temperature range between 4,900 and 6,500 K. These are unevolved, slowly rotating non-active stars from a CORALIE catalogue[11]. The stars have been monitored with high-precision spectroscopic observations for years to detect planetary systems. Of these 451 stars, 70 are reported to host planets and the rest, which we designate as a comparison sample (we often call them 'single' stars), have no detected planets so far. If there are planets around these single stars, their masses and orbital parameters will be different from those already known. We use this comparison sample to show that the reason for this extra Li depletion is not related to high metallicity (characteristic of planet-hosting stars) or to age (old stars are more Li-depleted).

Our abundance analysis, which followed standard prescriptions for stellar models, spectral synthesis code and stellar parameter determination[12], confirm the peculiar behaviour of Li in the effective temperature range 5,600–5,900 K for the 30 planet-bearing stars with respect to the 103 stars without planets in the comparison sample. To put this on a more solid statistical foundation, these two samples in the $T_{\text{eff}} = 5{,}600$–$5{,}900$ K window were extended by adding 16 planet-hosting and 13 comparison-sample stars, for which we have obtained new Li abundances from high-quality spectroscopic observations using the same spectral synthesis tools. We found that the immense majority of planet-hosting stars have severely depleted Li, whereas in the comparison sample a large fraction has only partially inhibited depletion. At higher and lower temperatures planet-hosting stars do not appear to show any peculiar behaviour in their Li abundance. The explanation of Li survival at $T_{\text{eff}} \gtrsim 5{,}850$ K is that the convective

layers of stars more massive than the Sun are shallow and too remote to reach the Li-burning layers. However, lower-mass stars with $T_{\text{eff}} \lesssim 5{,}700$ K have deeper convective layers that transport surface material to high-temperature regions in their interiors where Li can be destroyed more efficiently.

The Li abundance of some 20% of stars with exoplanets in the temperature range 5,600–5,900 K is $\log[N(\text{Li})] \geq 1.5$ (in standard notation, $\log[N(\text{Li})] = \log[n(\text{Li})/n(\text{H})] + 12$, where $n$ is the number density of atoms), while for the 116 comparison stars the Li abundance shows a rather high dispersion, with some 43% of the stars displaying Li abundances of $\log[N(\text{Li})] \geq 1.5$. This result becomes more obvious in solar-analogue stars, for which some 50% of the 60 single stars in the narrow window of $T_{\text{Sun}} \pm 80$ K (for $T_{\text{Sun}} = 5{,}777$ K) have $\log[N(\text{Li})] \geq 1.5$, while only two planet-hosting stars out of 24 have $\log[N(\text{Li})] \geq 1.5$ (Fig. 1). We performed different two-sample statistical tests using ASURV[13] (version 1.2). All tests consistently confirm
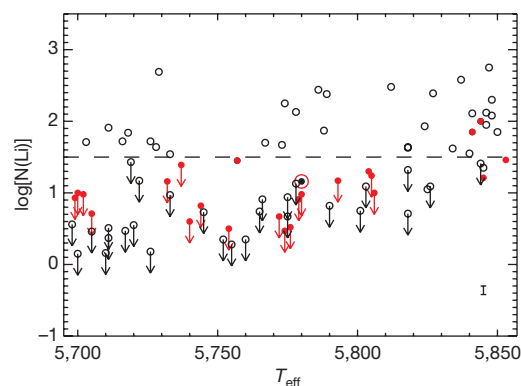


**Figure 1 | Lithium abundance plotted against effective temperature in solar-analogue stars with and without detected planets.** The planet-hosting and single stars are shown by filled red and empty black circles, respectively. The red circle with the black point at its centre indicates the Sun. The minimum detectable Li abundance varies among the stars used in this study because their spectra have different signal-to-noise ratios. The straight line at $\log[N(\text{Li})] = 1.5$ matches the upper envelope of the lower limits corresponding to a minimum signal-to-noise ratio of 200 in a typical solar twin. We use this line as a cut-off for selecting Li-depleted stars in our sample. We note that the two planet-hosting stars with the highest Li abundance also have nearly the highest effective temperatures and therefore thinner convective zones, which help to preserve Li. Other than in these stars, $\log[N(\text{Li})] = 1.5$ is the highest value found in a planet-hosting star. The mean statistical errors (1σ) for $\log[N(\text{Li})]$ and $T_{\text{eff}}$ averaged over all stars are 0.06 dex and 30 K, respectively[12]. Errors in $\log[N(\text{Li})]$ (bar in bottom right corner) include uncertainties in $T_{\text{eff}}$ and equivalent width measurement.

[1]Instituto de Astrofísica de Canarias, Via Láctea s/n, E-38200 La Laguna, [2]Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain. [3]Centro de Astrofísica, Universidade de Porto, Rua das Estrelas, 4150-762 Porto, Portugal. [4]Observatoire de Genève, Université de Genève, 51 ch des Maillettes, CH-1290 Versoix, Swtzerland. [5]Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain. [6]Instituto Nazionale di Astrofisica, Osservatorio di Arcetri, Largo Fermi 5, I59125 Firenze, Italy.

(at the $3\sigma$ level) that the planet-hosting and single star populations are not drawn from the same parent population. Subgiants were not included in this study because they undergo dramatic changes in their internal structure that alter the surface abundance of Li. The Li over-depletion in planet-bearing main-sequence stars is a generic feature over the $T_{eff}$-restricted range $T_{Sun} \pm 80$ K and is independent of $T_{eff}$ (or mass). These stars have very similar masses and similar surface convective zone depth, so there should be additional reasons for the over-depletion of Li. We now discuss the impact of age and metallicity on the Li abundance of solar-analogue stars.

The Li abundance of solar-type stars is expected to decrease progressively with age[14,15]. It is in principle possible that solar-analogue planet-hosting stars are on average older than the comparison sample and have depleted more Li. If that were the case, we should also expect a correlation between Li and stellar age indicators. Chromospheric activity is a reliable age indicator for solar-type stars from young ages up to about[15,16] 1 Gyr, or perhaps even to the age of the Sun[17]. Abundances of Li versus chromospheric activity indices[17], $R_{HK}$, are shown in Fig. 2a for the solar-analogue stars with and without detected planets. The comparison of the $R_{HK}$ values for the stars in our sample and for stars in the 625-Myr-old[15] Hyades cluster[18] indicates a much older age for our stars. We find no correlation between Li and the activity index in both samples (Fig. 2a). This suggests that age is not the main parameter governing Li depletion in our targets. It is known[19] that chromospheric activity correlates with stellar rotation ($v\sin i$). If the planet-hosting stars were older than the comparison-sample stars, their rotational velocities would be smaller than in the comparison sample. This is not observed either (Fig. 2b), adding support to our previous conclusion.

Most of the planet-hosting stars discovered so far are metal-rich[20]. The metallicity excess could result from either the accretion of planets/planetesimals onto the star or the protostellar molecular cloud. This metallicity excess is also present in the solar-analogue planet-hosting stars (see Fig. 2c). Can high metallicity be responsible for enhanced Li depletion in these stars? The increase of metal opacities in solar-type stars is responsible for the transition between radiative and convective energy transport. The main contributors to the total opacity at the base of the convective zone are oxygen and iron[21]. Our data (Fig. 2c) show that the fraction of single stars with $\log[N(Li)] > 1.5$ is 50% at both [Fe/H] $< 0$ and [Fe/H] $> 0$. This suggests that the Li-depletion mechanism does not depend on the metallicity in the range $-0.5 <$ [Fe/H] $< +0.5$. We have investigated the dependence of $\log[N(Li)]$ on [O/Fe] for planet-hosting stars, using oxygen data from the literature[22], and again found no correlation.

Comparison with field stars then leads us to the conclusion that neither age nor metallicity is responsible for the excess Li depletion. This is reinforced by observations of Li in solar-type stars in old clusters, which indeed show a wide dispersion of Li abundances with values ranging from $\log[N(Li)] = 2.5$ down to 1.0 and lower[5,23]. This is the case for M67 (age 3.5–4.8 Gyr and [Fe/H] $= 0.06$)[23] and NGC 6253 (age 3 Gyr and [Fe/H] $= 0.35$)[24,25], as is clearly seen in Fig. 2d. These two clusters offer a homogeneous sample of solar analogues in terms of age and metallicity. Solar analogues with both high and low Li abundance are present in these two clusters. The high Li abundance in a large fraction of old metal-rich stars in NGC 6253 and M67 leads us to conclude that high metallicity and/or age may not be the main cause for the systematic low Li abundances in solar-analogue planet-hosting stars. Our observations do not suggest that
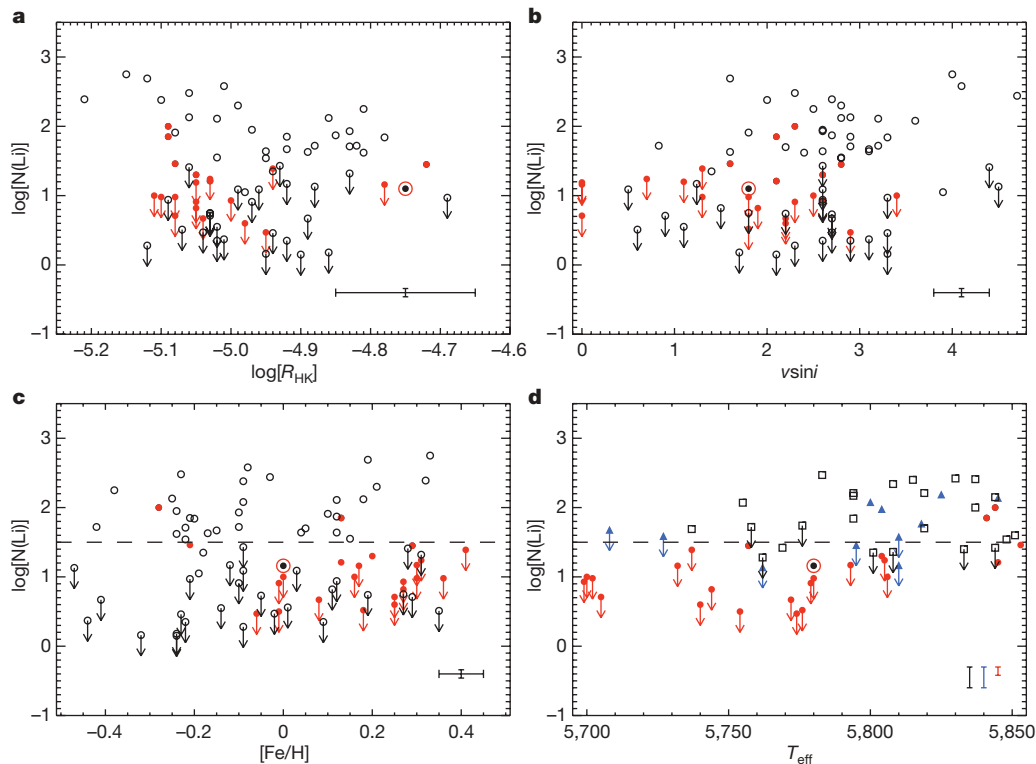


**Figure 2 | Lithium and physical properties of planet-hosting solar-analogue and comparison stars.** Chromospheric activity indices $R_{HK}$ were taken from the literature[17,27,28], while rotational velocities of the comparison-sample stars and many planet-hosting stars were measured from CORALIE and HARPS spectra using a cross-correlation function[29]. Typical $1\sigma$ uncertainties for $\log[R_{HK}]$ and $v\sin i$ (bars in bottom right corners of **a** and **b**) are 0.1 dex and 0.3 km s$^{-1}$, respectively[17,29]. Rotational velocities of several planet-hosts were taken from the literature[29,30]. The metallicities were measured[12] with a $1\sigma$ precision of 0.05 dex (bar in bottom right corner of **c**).

In **c** we plot Li abundances versus effective temperature in planet-hosting stars (filled red circles), and stars of the open clusters M67 (filled blue triangles) and NGC 6253 (empty black squares). The data for M67 were taken from the literature[15]. Li abundances in NGC 6253 have been derived from the Giraffe spectrograph attached to the Very Large Telescope using standard methods (S.R. *et al.*, manuscript in preparation). Typical $1\sigma$ error bars (bars in bottom right corner of **d**) are 0.15 dex and 100 K for $\log[N(Li)]$ and $T_{eff}$, respectively.

Li is unaffected by metallicity and/or age. They only imply that these parameters are not important enough to produce the enhanced Li depletion that we observe in solar-analogue stars with exoplanets.

We propose that the low Li abundance of planet-hosting solar-analogue stars is directly associated with the presence of planets. The presence of a planetary system may affect the angular momentum evolution of the star and the surface convective mixing. Planet migration will probably increase the angular momentum of the star. Various theoretical studies[3,6,7] show how magnetic braking scales with rotational velocity, leading to turbulent diffusion mixing and enhanced Li depletion. If that were the case we would expect severely Li-depleted stars to host planets with shorter orbital periods. There is no indication for such a correlation in the data, but we also note that in most cases we can only impose upper limits on the Li abundance, so that such correlations with orbital parameters could still be masked in the current data.

Alternatively, a long-lasting star–disk interaction during the pre-main sequence may cause planet-hosting stars to be slow rotators and develop a high degree of differential rotation between the radiative core and the convective envelope, also leading to enhanced Li depletion[26]. Revealing the relationships between protoplanetary disks and stellar structure in the early phases of the evolution of solar-type stars is a challenge for evolutionary models and simulations. It is possible that the enhanced Li depletion already takes place in the pre-main-sequence stage of planet-hosting stars. Exoplanet searches in very young stars will be needed to elucidate this. Asteroseismological observations of solar twins with and without known planets may reveal peculiarities in the inner structure of planet-hosting stars that could be the key to ascertaining the impact of planetary systems on the structure and angular momentum history of these stars.

It is known that solar-type stars with high metallicity have a high probability of hosting planets. Those solar analogues with low Li content (which is extremely easy to detect with simple spectroscopy) have an even higher probability of hosting exoplanets. Understanding the long-lasting mystery of the low Li abundance in the Sun appears to require proper modelling of the impact of planetary systems on the early evolution of solar-analogue stars.

1. Anders, E. & Grevesse, N. Abundances of the elements—meteoritic and solar. *Geochim. Cosmochim. Acta* **53**, 197–214 (1989).
2. D'Antona, F. & Mazzitelli, I. New pre-main-sequence tracks for M less than or equal to 2.5 solar mass as tests of opacities and convection model. *Astrophys. J. Suppl. Ser.* **90**, 467–500 (1994).
3. Maeder, A. *Physics, Formation and Evolution of Rotating Stars* 215–351 (Astron. Astrophys. Library, Springer, 2009).
4. Favata, F., Micela, G. & Sciortino, S. Lithium abundance in a volume-limited sample of nearby main sequence G and K stars. *Astron. Astrophys.* **311**, 951–960 (1996).
5. Randich, S. Light element abundances in solar-type members of open clusters. *Mem. Soc. Astron. Ital.* **79**, 516–523 (2008).
6. Pinsonneault, M. Mixing in stars. *Annu. Rev. Astron. Astrophys.* **35**, 557–605 (1997).
7. Pinsonneault, M., Kawaler, S. D., Sofia, S. & Demarque, P. Evolutionary models of the rotating sun. *Astrophys. J.* **338**, 424–452 (1989).
8. Israelian, G., Santos, N., Mayor, M. & Rebolo, R. Lithium in stars with exoplanets. *Astron. Astrophys.* **414**, 601–611 (2004).
9. Gonzalez, G. Parent stars of extrasolar planets. IX. Lithium abundances. *Mon. Not. R. Astron. Soc.* **386**, 928–934 (2008).
10. Takeda, Y., Kawanomoto, S., Honda, S., Ando, H. & Sakurai, T. Behavior of Li abundances in solar-analog stars. Evidence for line-width dependence. *Astron. Astrophys.* **468**, 663–677 (2007).
11. Mayor, M. *et al.* Setting new standards with HARPS. *The Messenger* **114**, 20–24 (2003).
12. Sousa, S. *et al.* Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes. *Astron. Astrophys.* **487**, 373–381 (2008).
13. Feigelson, E. D. & Nelson, P. I. Statistical methods for astronomical data with upper limits. I—Univariate distributions. *Astrophys. J.* **293**, 192–206 (1985).
14. Zappala, R. R. Lithium abundances of stars in open clusters. *Astrophys. J.* **172**, 57–74 (1972).
15. Sestito, P. & Randich, S. Time scales of Li evolution: a homogeneous analysis of open clusters from ZAMS to late-MS. *Astron. Astrophys.* **442**, 615–627 (2005).
16. Pace, G. *et al.* An investigation of chromospheric activity spanning the Vaughan-Preston gap: impact on stellar ages. *Astron. Astrophys.* **499**, L9–L12 (2009).
17. Wright, J., Marcy, G. W., Butler, R. P. & Vogt, S. S. Chromospheric Ca II emission in nearby F, G, K, and M stars. *Astrophys. J. Suppl. Ser.* **152**, 261–295 (2004).
18. Paulson, D. B., Saar, S. H., Cochran, W. D. & Hatzes, A. P. Searching for planets in the Hyades. II. Some implications of stellar magnetic activity. *Astron. J.* **124**, 572–582 (2002).
19. Cutispoto, G. *et al.* Fast-rotating nearby solar-type stars sin i and X-ray luminosities relationships. II. Li abundances, v sin i and X-ray luminosities relationships. *Astron. Astrophys.* **397**, 987–995 (2003).
20. Santos, N. C., Israelian, G. & Mayor, M. Spectroscopic [Fe/H] for 98 extra-solar planet-host stars. Exploring the probability of planet formation. *Astron. Astrophys.* **415**, 1153–1166 (2004).
21. Piau, L. & Turck-Chièze, S. Lithium depletion in pre-main-sequence solar-like stars. *Astron. Astrophys.* **566**, 419–434 (2002).
22. Ecuvillon, A. *et al.* Oxygen abundances in planet-harbouring stars. Comparison of different abundance indicators. *Astron. Astrophys.* **445**, 633–645 (2006).
23. Pasquini, L., Biazzo, K., Bonifacio, P., Randich, S. & Bedin, L. R. Solar twins in M 67. *Astron. Astrophys.* **489**, 677–684 (2008).
24. Yadav, R. *et al.* Ground-based CCD astrometry with wide-field imagers. II. A star catalog for M 67: WFI@2.2 m MPG/ESO astrometry, FLAMES@VLT radial velocities, *Astron. Astrophys.* **484**, 609–620 (2008).
25. Sestito, P., Randich, S. & Bragaglia, A. Element abundances in the metal-rich open cluster NGC 6253. *Astron. Astrophys.* **465**, 185–196 (2007).
26. Bouvier, J. Lithium depletion and the rotational history of exoplanet host stars. *Astron. Astrophys.* **489**, L53–L86 (2008).
27. Saffe, C., Gómez, M. & Chavero, C. On the ages of exoplanet host stars. *Astron. Astrophys.* **443**, 609–626 (2005).
28. Gray, R. *et al.* Contributions to the nearby stars (NStars) project: spectroscopy of stars earlier than M0 within 40 pc—the southern sample. *Astron. J.* **132**, 161–170 (2006).
29. Santos, N. C. *et al.* The CORALIE survey for southern extra-solar planets. IX. A 1.3-day period brown dwarf disguised as a planet. *Astron. Astrophys.* **392**, 215–229 (2002).
30. Valenti, J. A. & Fischer, D. A. Spectroscopic properties of cool stars (SPOCS). I. 1040 F, G, and K dwarfs from Keck, Lick, and AAT planet search programs. *Astrophys. J. Suppl. Ser.* **159**, 141–166 (2005).

**Author Contributions** All authors participated in data collection, analysis, interpretation and commented on the manuscript. G.I. led the project and wrote the paper.

# LETTERS

# Fractional quantum Hall effect and insulating phase of Dirac electrons in graphene

Xu Du[1]†, Ivan Skachko[1], Fabian Duerr[1], Adina Luican[1] & Eva Y. Andrei[1]

In graphene, which is an atomic layer of crystalline carbon, two of the distinguishing properties of the material are the charge carriers' two-dimensional and relativistic character. The first experimental evidence of the two-dimensional nature of graphene came from the observation of a sequence of plateaus in measurements of its transport properties in the presence of an applied magnetic field[1,2]. These are signatures of the so-called integer quantum Hall effect. However, as a consequence of the relativistic character of the charge carriers, the integer quantum Hall effect observed in graphene is qualitatively different from its semiconductor analogue[3]. As a third distinguishing feature of graphene, it has been conjectured that interactions and correlations should be important in this material, but surprisingly, evidence of collective behaviour in graphene is lacking. In particular, the quintessential collective quantum behaviour in two dimensions, the fractional quantum Hall effect (FQHE), has so far resisted observation in graphene despite intense efforts and theoretical predictions of its existence[4–9]. Here we report the observation of the FQHE in graphene. Our observations are made possible by using suspended graphene devices probed by two-terminal charge transport measurements[10]. This allows us to isolate the sample from substrate-induced perturbations that usually obscure the effects of interactions in this system and to avoid effects of finite geometry. At low carrier density, we find a field-induced transition to an insulator that competes with the FQHE, allowing its observation only in the highest quality samples. We believe that these results will open the door to the physics of FQHE and other collective behaviour in graphene.

The description of graphene in terms of a two-dimensional (2D) zero-bandgap semiconductor with low energy excitations represented by non-interacting Dirac fermions is surprisingly successful[3]. Indeed, most experimental results thus far are captured by this single particle picture, in which collective effects and interactions are assumed to be negligibly small. In particular, scanning tunnelling spectroscopy in a transverse magnetic field, $B$, revealed a sequence of Landau levels with energy $E_n = \text{sign}(n)\sqrt{2e\hbar v_F^2|n|B}$, providing the most direct evidence of the non-interacting Dirac fermion picture[11]. (Here $n = 0,\pm 1,\pm 2$, $e$ is the elementary charge, $\hbar = h/2\pi$ where $h$ is Planck's constant and $v_F$ is the Fermi velocity[3]). Sweeping the field or carrier density, $n_s$, through these Landau levels in a magneto-transport measurement reveals quantum-Hall conductance plateaus (the integer quantum Hall effect) at values $G_{xy} = v\frac{e^2}{h}$ for filling factors $v = \frac{n_s}{B}\frac{h}{e} = 4(n+1/2)$ where all available states up to the $n$th Landau level are occupied. Here 4 is due to the spin and valley degeneracy[3] and the $\pm$ signs reflect the electron–hole symmetry. The $\frac{1}{2}\frac{4e^2}{h}$ offset, absent in non-relativistic 2D electron systems (2DES), is a result of the special status of the $n = 0$ Landau level for the massless Dirac fermions: half of its states are hole states, and the other half are electron states. This picture is expected to fail when interactions

lift the degeneracy, resulting in new integer plateaus outside this sequence[12–15]. Furthermore, strong correlations between the electrons are expected to give rise to plateaus at fractional filling factors, reflecting the condensation into new ground states[4–9].

Thus far, magneto-transport experiments on non-suspended graphene samples show no evidence of interactions or correlations for fields below ~25 T. In higher fields[16], the appearance of quantum Hall effect plateaus at $v = 0,\pm 1,\pm 4$ suggests that interaction effects do exist in graphene, but only become observable when their energy scale exceeds that of the fluctuations due to random potentials induced by external sources. Similarly, the insulating state at $v = 0$ observed in some non-suspended samples in strong magnetic fields[17] but not in others[18] suggests that extrinsic effects play an important role in obscuring the underlying intrinsic physics of the charge carriers in graphene. Therefore in order to understand the role of correlations in the low density phases and to solve the long-standing mystery of whether graphene can support an FQHE, it is necessary to better control sample quality.

Recently, a significant improvement in transport properties was demonstrated in suspended graphene samples where substrate-induced perturbations were eliminated[10,19]. The combination of ballistic transport and low carrier density achieved in suspended graphene is particularly well suited for studying the intrinsic properties of this system. To ensure mechanical and structural integrity of the sample, the suspended devices are quite small, with typical dimensions of length $L \approx 0.5$–1 μm and width $W \approx 1.5$–3 μm. Surprisingly, in these small devices the standard Hall-bar measurement geometry fails to yield the expected quantum Hall effect features[19]. A possible cause, which we discuss in detail elsewhere, is that the proximity between voltage and current leads in such small samples shorts out most of the Hall voltage. This is a consequence of the peculiar potential distribution at large Hall angles (the case of plateaus in the quantum Hall effect) where most of the potential drop, roughly equal to the Hall voltage, occurs at opposite corners of the sample close to the current leads[20], also known as hot spots in the Hall effect literature. A metallic lead placed within this region necessarily shorts out the Hall voltage. As suspended graphene devices (typically of micrometre size) are too small to allow the voltage leads to be placed outside the hot spot regions, shorting the Hall voltage is almost unavoidable[19]. This problem is circumvented in the two-terminal lead configuration[10] used in the present work (Fig. 1a). We note that the classical contact resistance of the two-terminal devices discussed here is negligible compared to the quantum resistance, as indicated by the very small deviation of the two-terminal resistance at the quantum Hall effect plateaus from the standard values, described below.

All the suspended graphene samples studied here are in the ballistic regime, as measured by the density dependence of the zero field conductivity. In the hole-carrier sector, we find the mean-free-path

[1]Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08855, USA. †Present address: Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA.
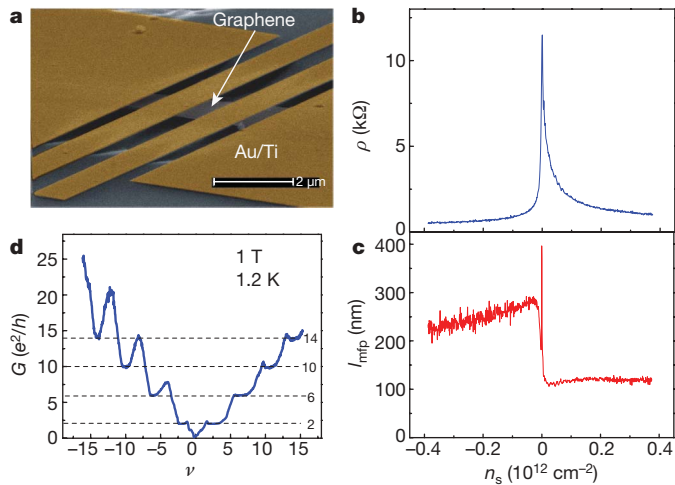
**Figure 1 | Characteristics of the suspended graphene devices. a,** False-colour scanning electron microscopy image of a typical suspended graphene device. The two centre pads are used for both current and voltage leads, while the outer pads are for structural support. The lead separation is $L = 0.7 \, \mu m$, and the typical graphene width is 1.5–3 $\mu m$. **b,** Carrier density dependence of the resistivity of a suspended graphene device in zero field. The sharp gate control of resistivity near the Dirac point indicates a low level of perturbation from random potentials. **c,** Carrier density dependence of the mean free path, $l_{mfp} = \frac{\sigma h}{2e^2 (\pi n_s)^{1/2}}$, of the sample in **b**. Note that on the hole branch, $l_{mfp} \approx L/2$, as expected for ballistic junctions. **d,** Conductance of the suspended graphene sample as a function of filling factor $v$ for $B = 1$ T and $T = 1.2$ K. The plateaus seen at integer filling factors correspond to the quantum Hall effect, as discussed in the text. The maxima in between the plateaus agree with the theoretical expectations[21] for a two-terminal graphene junction with the geometry of our sample, $W/L > 1$. The quantum Hall plateaus are better defined and narrower for the hole branch (negative filling factors), indicating less scattering of hole carriers, consistent with the lower resistance and longer mean free path on the hole branch, as shown in **b** and **c**.

$l_{mfp} \approx L/2$ and the conductivity $\sigma \propto n_s^{1/2}$, as expected for ballistic transport (Fig. 1b, c). Furthermore the lowest carrier density, typically $n_{s0} \approx (2–10) \times 10^9 \, cm^{-2}$, is more than an order of magnitude below that achieved in non-suspended samples, attesting to a much smaller density inhomogeneity[10]. For non-ballistic samples (graphene as well as 2DES in semiconductors), the sample quality is usually characterized by the carrier mobility. In ballistic graphene samples however, the value of mobility is meaningful only when it is associated with the carrier density at which it is measured. For the sample studied here, the Drude mobility, $\mu_D = \sigma/n_s e$, at $n_s \approx 10^{10} \, cm^{-2}$ is $260{,}000 \, cm^2 \, V^{-1} \, s^{-1}$, and exhibits the $\propto n_s^{-1/2}$ dependence on carrier density expected for ballistic devices (the field effect mobility at the same density is $\mu_{fe} = \frac{1}{e} \frac{d\sigma}{dn_s} \approx 200{,}000 \, cm^2 V^{-1} s^{-1}$).

We studied the two-terminal magneto-transport in suspended graphene samples at temperatures ranging from 1.2 K to 80 K and fields up to 12 T. The relation between magneto-resistance oscillations and the quantum Hall effect measured in two-terminal devices is now well understood. It has been shown theoretically[21] that, for clean samples and low temperatures, the two-terminal conductance displays plateaus at values $G = v \frac{e^2}{h}$ that are precisely the same as the quantum Hall effect plateaus in the Hall conductance. In between the plateaus the conductance is non-monotonic, depending on the sample aspect ratio, $W/L$. In our devices where $W > L$, the conductance is expected to overshoot between plateaus, as is indeed observed (Fig. 1d). Our two-terminal measurements reveal well-defined plateaus associated with the anomalous quantum Hall effect that appear already in fields below 1 T. Above 2 T additional plateaus develop at $v = -1$ and at $v = 3$, reflecting interaction-induced lifting of the spin and valley degeneracy (Figs 2a and 3c). At low temperatures and above 2 T, we observe a FQHE plateau at $v = -1/3$ which becomes better defined with increasing field (Fig. 2a). When plotting $G$ versus $v$, the curves for all values of $B$ collapse together (Fig. 2b), and the plateaus at $v = -1/3$, $-1$ and $-2$ show accurate values of the quantum Hall conductance.

The FQHE in semiconductor based 2DES reflects the formation of an incompressible condensate, which can be described by a Laughlin wavefunction[22]. In the composite-fermion generalization of the FQHE[4,23], a strongly correlated electron liquid in a magnetic field can minimize its energy when the filling factor belongs to the series $v = \frac{p}{2sp \pm 1}$ (with $s$ and $p$ integers) by forming weakly interacting composite particles consisting of an electron and an even number of captured magnetic flux lines. In this picture, the FQHE with $v = 1/3$ corresponds to the integer quantum Hall effect with $v = 1$ for the composite particles consisting of one electron and two flux lines. Excitations out of this state would produce fractionally charged quasiparticles $q^* = e/3$, at an energy cost of the excitation gap, $\Delta_{1/3}$, which provides a measure of the state's robustness. It is not obvious a priori that the correlated state leading to the FQHE for the relativistic charge carriers in graphene is the same as that for the 2DES in semiconductors. In fact, several competing mechanisms have been discussed in the theoretical literature[4–9], involving states that break SU(4) symmetry as well as possible compressible, composite fermion Fermi sea states[7]. Interestingly, despite the qualitative difference in Landau level spectra between Dirac fermions in graphene and the non-relativistic electrons in semiconductors, the $v = 1/3$ state is formally expected to be the same in both cases[4,5] but with the pseudospin in graphene playing the role of the traditional electron spin in the non-relativistic case. In order to distinguish experimentally between the various mechanisms, it is useful to study the quasiparticle excitation energy. In multi-lead transport measurements, such as the Hall bar configuration, this can be obtained from the temperature dependence of the longitudinal conductance. However, in a two-terminal measurement it is not possible to separate the longitudinal



**Figure 2 | FQHE in suspended graphene. a,** Gate voltage dependence of resistance for the sample in Fig. 1, at indicated magnetic fields and $T = 1.2$ K. Already at 2 T we note the appearance of quantum Hall plateaus outside the non-interacting sequence, with $R = \frac{1}{v} \frac{h}{e^2}$, $v = 1, 1/3$. **b,** Hole conductance as a function of filling factors for $B = 6$ T, 8T, 10T and 12T at $T = 1.2$ K, showing

that the data collapse together. Quantum Hall plateaus with conductance values $G = v \frac{e^2}{h}$, $v = 1, 1/3$, appear at the correct filling factors of $v = -1$, $-1/3$. **c,** Temperature dependence of the quantum Hall plateau features. The plateaus at $v = -1/3$, $-1$ become smeared out with increasing $T$ and disappear for $T > 20$ K.
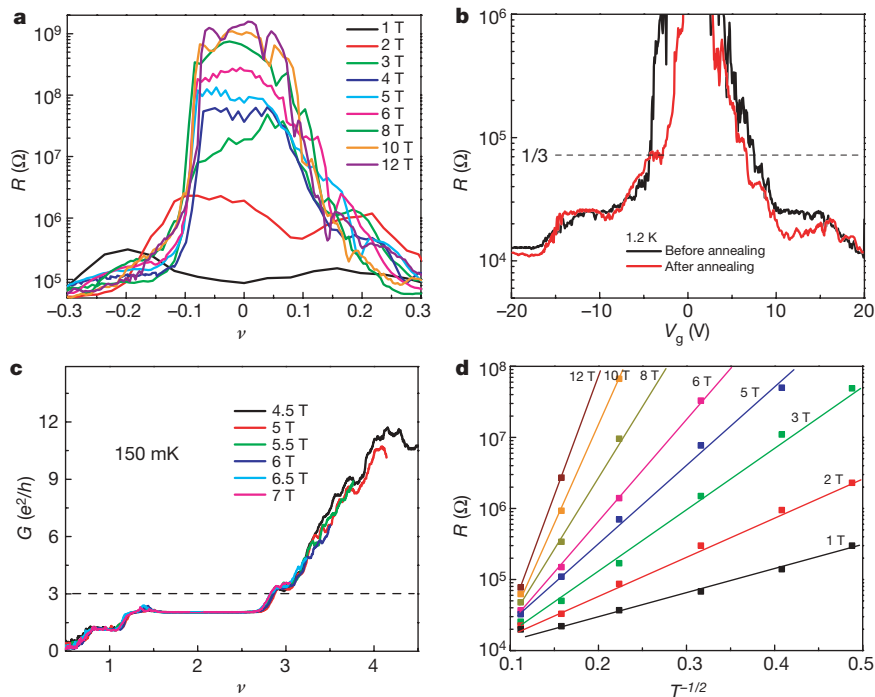
**Figure 3 | Insulating behaviour at $v = 0$. a**, Resistance as a function of filling factor for magnetic fields $B = 1,2,3,4,5,6,8,10,12$ T. For $|v| < 0.1$, the resistance increases sharply with increasing magnetic field. The maximum resistance value measured above 8 T is instrument-limited. **b**, Competition between FQHE and insulating behaviour. The sample in Fig. 1 was warmed up to room temperature and re-cooled to 1.2 K. Owing to the condensation of contaminants on the graphene channel, the insulating regime became broader, swallowing the FQHE plateau at $v = -1/3$. On current annealing, the sample was re-cleaned almost to its pristine condition, causing the insulating regime to recede and the plateau at $v = -1/3$ to reappear. **c**, Quantum Hall effect plateaus of a suspended graphene sample which showed $v = 3$. **d**, Logarithmic plot of maximum resistance for $v = 0$ as a function of $T^{-1/2}$ for the field values shown in **a**. The solid lines are guides to the eye.

and transverse components. Nevertheless, an order of magnitude estimate can be obtained from the temperature at which the $v = 1/3$ plateau disappears. In Fig. 2c we note that this plateau smears out with increasing temperature and disappears above 20 K, suggesting that $\Delta_{1/3} \approx 20\,\text{K} \approx 0.008E_C(12\text{T})$, where $E_C = e^2/4\pi\varepsilon_0\varepsilon\,l_B = 650B^{1/2}\,\text{K}$ (with $B$ in units of tesla) is the Coulomb energy, $\varepsilon_0$ the permittivity of free space, $\varepsilon = 1$ the dielectric constant of the host material (vacuum) and $l_B = \sqrt{\frac{\hbar}{eB}}$ is the magnetic length. This is about 8% of the theoretical prediction[4,5], $\Delta_{1/3} \approx 0.1E_C$, for the gap in a Laughlin-like condensate. For 2DES in semiconductors, the discrepancy with theoretical predictions is larger, reflecting deviations from an ideal 2D system due to the finite thickness of the quantum wells (10–30 nm), disorder and mixing with higher Landau levels[24]. Importantly, the value of $\Delta_{1/3}$ in suspended graphene is more than an order of magnitude larger than the corresponding gap in the 2DES in semiconductors[25] because of the lower dielectric constant ($\varepsilon = 1$ in suspended graphene compared to $\varepsilon \approx 12.9$ in GaAs/GaAlAs heterostructures).

Next we discuss transport near the Dirac point ($v = 0$). Models for lifting of the fourfold spin and valley degeneracy fall in two categories, depending on whether the spin degeneracy is lifted first, producing a so-called quantum Hall ferromagnet[12–14] or the valley degeneracy is lifted first, which gives rise to magnetic catalysis[14,15]. Both cases predict insulating bulk, but the former supports counter-propagating edge states and thus is a conductor, whereas the latter with no edge states is an insulator. In the quantum Hall ferromagnet scenario, where both spin and valley degeneracy can be lifted for all Landau levels, plateaus at all integer $v$ are allowed. In contrast, the magnetic catalysis scenario does not permit plateaus at odd filling-factors other than $v = \pm1$. Experiments addressing this issue in non-suspended graphene are inconclusive[16–18]. While tilted field experiments support the quantum Hall ferromagnet situation[16], the absence of clear plateaus at $\pm3,\pm5$ is consistent with magnetic catalysis. The fact that both insulating and conducting behaviour were reported further contributes to the uncertainty.

To address this question in suspended graphene samples, we studied four samples in fields up to 12 T and at temperatures ranging from 1 K to 80 K (Supplementary Information). All samples were insulating at $v = 0$ for high fields and low temperature. Consistently we found that the higher the sample quality, as measured by the residual carrier density, the sharper the transition to the insulating state and the earlier its onset (lower fields and higher temperatures). In our highest quality sample, the onset of insulating behaviour scales linearly with field. This is clearly seen in Fig. 3a, where the sharp onset of insulating behaviour at $|v| \approx 0.1$ is marked by a dramatic increase in resistance. In the best samples, the maximum resistance value is instrument-limited to $\sim$1 G$\Omega$. In lower quality samples, the insulating region is broader, the onset less sharp and the maximum resistance lower. Interestingly, the FQHE state was only observed in samples with narrow insulating regions, suggesting a competition between the two ground states. This is illustrated in Fig. 3b, where the insulating phase, having become broader after contamination, 'swallowed' the 1/3 plateau. Current annealing the sample brought it back almost to its pristine condition again, revealing the 1/3 plateau.

Can the suspended graphene data shed light on the nature of the insulating phase? The appearance of a plateau at $v = 3$, shown in Fig. 3c, favours the quantum Hall ferromagnet over the magnetic catalysis. However, since the quantum Hall ferromagnet supports counter-propagating edge states, this scenario is inconsistent with insulating behaviour at $v = 0$. A possible solution would entail a gap opening in the edge states and thus a mechanism to admix them. This would require a mechanism to flip spins and valleys, such as magnetic impurities or segments of zigzag edges[26]. An alternative explanation is that the system undergoes a transition to a new broken symmetry phase, such as a Wigner crystal or a more exotic skyrme phase[27,28]. In this case pinning would naturally lead to insulating behaviour.

To better understand the insulating phase, we studied the temperature dependence of the $v = 0$ state. The details of the temperature dependence of the maximum resistance ($R_{max}$) show strong sample-to-sample variation, but all curves fit a generalized activated form: $R_{max} = R_0 \exp(T_0/T)^\alpha$ with $\alpha \approx 1/3 - 1$. In the best sample (Fig. 3c), $\alpha \approx 1/2$ for all fields, with $T_0 \propto B^2$. This may provide a hint to the nature of the insulating state, but more work is needed to resolve this question.

In summary, the experiments described here demonstrate that Dirac electrons do exhibit strong collective behaviour leading to an FQHE, which becomes apparent in suspended samples probed with a two-terminal lead geometry, where the system is isolated from external perturbations. We find that the FQHE is quite robust, appearing at low temperatures in fields as low as 2 T and persisting up to 20 K in a field of 12 T. The effect is significantly more robust than in the semiconductor-based 2DES, reflecting the stronger Coulomb interaction and the more 2D nature of the 2DES in graphene. We further show that the FQHE state competes with an insulating phase centred at $v = 0$ that broadens in the presence of disorder and can destroy it. This may explain why, despite the large energy scale of the Coulomb interactions, the FQHE has until now resisted observation in graphene. The observation of the FQHE plateau at $v = -1/3$ demonstrates that the FQHE is a stable ground state for the 2D Dirac fermions in graphene, and that it is a distinctly different phase from the insulating state at $v = 0$. These findings pave the way to future studies of FQHE physics in the Dirac fermion system for the $n = 0$ Landau level as well for higher Landau levels, where new correlated states, unique to relativistic charge carriers, are expected to emerge.

1.  Novoselov, K. S. et al. Two-dimensional gas of massless Dirac fermions in graphene. Nature 438, 197–200 (2005).
2.  Zhang, Y. B., Tan, Y. W., Stormer, H. L. & Kim, P. Experimental observation of the quantum Hall effect and Berry's phase in graphene. Nature 438, 201–204 (2005).
3.  Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. Rev. Mod. Phys. 81, 109–162 (2009).
4.  Toke, C., Lammert, P. E., Jain, J. K. & Crespi, V. H. Fractional quantum Hall effect in graphene. Phys. Rev. B 74, 235417 (2006).
5.  Yang, K., Das Sarma, S. & MacDonald, A. H. Collective modes and skyrmion excitations in graphene SU(4) quantum Hall ferromagnets. Phys. Rev. B 74, 075423 (2006).
6.  Peres, N. M. R., Guinea, F. & Castro Neto, A. H. Electronic properties of disordered two-dimensional carbon. Phys. Rev. B 73, 125411 (2006).
7.  Khveshchenko, D. V. Composite Dirac fermions in graphene. Phys. Rev. B 75, 153405 (2007).
8.  Shibata, N. & Nomura, K. Coupled charge and valley excitations in graphene quantum Hall ferromagnets. Phys. Rev. B 77, 235426 (2008).
9.  Goerbig, M. O. & Regnault, N. Analysis of a SU(4) generalization of Halperin's wave function as an approach towards a SU(4) fractional quantum Hall effect in graphene sheets. Phys. Rev. B 75, 241405 (2007).
10. Du, X., Skachko, I., Barker, A. & Andrei, E. Y. Approaching ballistic transport in suspended graphene. Nature Nanotechnol. 3, 491 (2008).
11. Li, G., Luican, A. & Andrei, E. Y. Scanning tunnelling spectroscopy of graphene. Phys. Rev. Lett. 102, 176804 (2009).
12. Nomura, K. & MacDonald, A. H. Quantum Hall ferromagnetism in graphene. Phys. Rev. Lett. 96, 256602 (2006).
13. Alicea, J. & Fisher, M. P. A. Integer quantum Hall effect in the ferromagnetic and paramagnetic regimes. Phys. Rev. B 74, 075422 (2006).
14. Yang, K. Spontaneous symmetry breaking and quantum Hall effect in graphene. Solid State Commun. 143, 27–32 (2007).
15. Gorbar, E. V., Gusynin, V. P., Miransky, V. A. & Shovkovy, I. A. Dynamics in the quantum Hall effect and the phase diagram of graphene. Phys. Rev. B 78, 085437 (2008).
16. Zhang, Y. et al. Landau level splitting in graphene in high magnetic fields. Phys. Rev. Lett. 96, 136806 (2006).
17. Checkelsky, J. G., Li, L. & Ong, N. P. Divergent resistance at the Dirac point in graphene: evidence for a transition in high magnetic field. Phys. Rev. Lett. 100, 206801 (2008).
18. Abanin, D. A. et al. Dissipative quantum Hall effect in graphene near the Dirac point. Phys. Rev. Lett. 98, 196806 (2007).
19. Bolotin, K. I. et al. Ultrahigh electron mobility in suspended graphene. Solid State Commun. 146, 351–355 (2008).
20. Wakabayashi, J. & Kawaji, S. Hall effect in silicon MOS inversion layers under strong magnetic fields. J. Phys. Soc. Jpn 44, 1839–1849 (1978).
21. Abanin, D. A. & Levitov, L. S. Conformal invariance and shape-dependent conductance of graphene samples. Phys. Rev. B 78, 035416 (2008).
22. Laughlin, R. B. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. Phys. Rev. Lett. 50, 1395–1398 (1983).
23. Jain, J. K. Composite fermion approach for the fractional quantum Hall effect. Phys. Rev. Lett. 63, 199–202 (1989).
24. Xin, W. et al. Mobility gap in fractional quantum Hall liquids: effects of disorder and layer thickness. Phys. Rev. B 72, 075325 (2005).
25. Boebinger, G. S. et al. Activation energies and localization in the fractional quantum Hall effect. Phys. Rev. B 36, 7919–7929 (1987).
26. Shimshoni, E., Fertig, H. A. & Pai, G. V. Onset of an insulating zero-plateau quantum hall state in graphene. Phys. Rev. Lett. 102, 206408 (2009).
27. Cote, R., Jobidon, J. F. & Fertig, H. A. Skyrme and Wigner crystals in graphene. Phys. Rev. B 78, 085309 (2008).
28. Poplavskyy, O., Goerbig, M. O. & Morais Smith, C. Local density of states of electron-crystal phases in graphene in the quantum Hall regime. Preprint at ⟨http://arXiv.org/abs/0902.1902⟩ (2009).

# LETTERS

# Observation of the fractional quantum Hall effect in graphene

Kirill I. Bolotin[1]*†, Fereshte Ghahari[1]*, Michael D. Shulman[2], Horst L. Stormer[1,2] & Philip Kim[1,2]

**When electrons are confined in two dimensions and subject to strong magnetic fields, the Coulomb interactions between them can become very strong, leading to the formation of correlated states of matter, such as the fractional quantum Hall liquid[1,2]. In this strong quantum regime, electrons and magnetic flux quanta bind to form complex composite quasiparticles with fractional electronic charge; these are manifest in transport measurements of the Hall conductivity as rational fractions of the elementary conductance quantum. The experimental discovery of an anomalous integer quantum Hall effect in graphene has enabled the study of a correlated two-dimensional electronic system, in which the interacting electrons behave like massless chiral fermions[3,4]. However, owing to the prevailing disorder, graphene has so far exhibited only weak signatures of correlated electron phenomena[5,6], despite intense experimental and theoretical efforts[7–14]. Here we report the observation of the fractional quantum Hall effect in ultraclean, suspended graphene. In addition, we show that at low carrier density graphene becomes an insulator with a magnetic-field-tunable energy gap. These newly discovered quantum states offer the opportunity to study correlated Dirac fermions in graphene in the presence of large magnetic fields.**

In a perpendicular magnetic field, the energy spectrum of a clean two-dimensional electron system (2DES) splits into a fan of Landau levels. When the Fermi energy is tuned to lie between the Landau levels, the system enters the integer quantum Hall regime, in which current is carried by states at the edge of the sample and the overall conductance ($G$) is quantized as $G = \nu e^2/h$, where $\nu$ is a Landau level filling factor which takes on integer values in the case of the integer quantum Hall effect (IQHE), $e$ is the elementary charge, and $h$ is Planck's constant. The first observation of the IQHE in graphene demonstrated an unusual sequence of filling factors, $\nu = \pm 2, \pm 6, \pm 10, ...$ which differs from previously studied 2DESs[2]. This sequence originates from two peculiar features of graphene: the fourfold spin and pseudo-spin (valley) degeneracy of Landau levels and the existence of a non-trivial Berry phase associated with the pseudo-spin of Dirac quasiparticles[15].

In clean samples and under very strong magnetic fields, additional integer quantum Hall states emerge at filling factors $\nu = 0, \pm 1$ (refs 5, 6). These fragile states are conjectured to result from electron–electron (e–e) interactions lifting the pseudospin and spin degeneracy of the zeroth Landau level[16]. The nature of these states, and in particular the unusual $\nu = 0$ state, responsible for the divergent resistivity of graphene at high magnetic fields[17], has raised considerable interest in the effect of e–e interactions among Dirac quasiparticles. In conventional semiconductor heterojunctions, such correlation effects are spectacularly manifested in the fractional quantum Hall effect (FQHE), where new electronic ground states are formed in which the elementary excitations are composite particles with fractional charge[18]. The possibility of the

FQHE in graphene and the interplay between many-body correlations and the SU(4) symmetry of quasiparticles have ignited considerable theoretical interest[7–14]. Thus far, however, strong electron scattering off omnipresent residual impurities have precluded the observation of the FQHE.

It has been previously demonstrated[19–21] that the scattering can be drastically reduced by fabricating suspended graphene devices. In such devices, the carrier mobility $\mu = (ne\rho)^{-1}$, where $n$ is the carrier density and $\rho$ is the resistivity, can exceed 200,000 cm$^2$ V$^{-1}$ s$^{-1}$. It was suggested that samples of this quality should be appropriate for the observation of the FQHE[22]. In the present study, we use two-terminal suspended graphene devices (Fig. 1a; see Supplementary Information
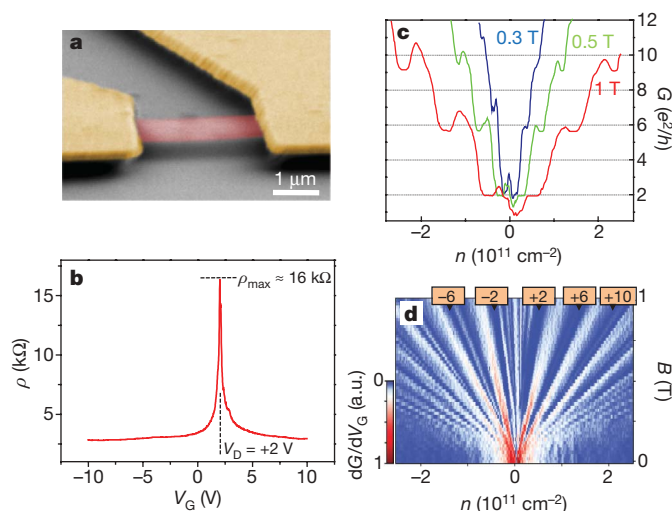


**Figure 1 | Electrical properties of suspended graphene at low magnetic field. a,** False-colour scanning electron micrograph of a typical device. Single layer graphene (red) is suspended ~150 nm above SiO$_2$/Si substrate (grey) and supported by two gold electrodes (yellow). To ensure mechanical stability of the devices, we use a two-terminal configuration in this study. **b,** Resistivity $\rho$ of graphene as a function of gate voltage $V_G$ applied between the graphene and the substrate, measured at $T = 2$ K and $B = 0$ T. The resistivity is calculated from the resistance $R$ via $\rho = RW/L$, where $W$ and $L$ are sample width and length, respectively. **c,** Conductance $G$ as a function of carrier density $n$ at different magnetic fields $B = 0.3, 0.5, 1$ T. **d,** Colour rendition of abs[d$G$/d$n(B,n)$] in the low magnetic field regime $B = 0$–1 T. Dark colour corresponds to quantum Hall plateau region $G(n) = \nu e^2/h = $ const. The gate capacitance $C_g \approx 50$ aF μm$^{-2}$ used in calculating the carrier density $n = C_g(V_G - V_D)$ is obtained from the shift of the quantum Hall plateaus with magnetic field. The contact resistance ~200 Ω is estimated from the deviation of the conductance from the expected quantized value $G = \nu e^2/h$.

[1]Department of Physics, [2]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York 10027, USA. †Present address: Department of Physics, Vanderbilt University, Nashville, Tennessee 37235, USA.
*These authors contributed equally to this work.

for fabrication details). These devices are mechanically more robust and tend to better survive thermal cycling, compared to the previously studied multiprobe specimens[19]. The devices are measured in a cryostat capable of magnetic fields $B$ up to 14 T and temperatures $T$ between 2 and 15 K. The two-terminal conductance $G$ is recorded as a function of $B$, $T$ and $n$, which is tuned by adjusting the voltage $V_G$ between the graphene and the back gate electrode. We limit $|V_G| < 10$ V in order to avoid collapse of the graphene due to electrostatic attraction, limiting the accessible carrier densities to $|n| < 3.5 \times 10^{11}$ cm$^{-2}$. At $T = 2$ K, a sharp peak in resistance $R = 1/G$ is evident around the Dirac point $V_D \approx 2$ V; at this gate voltage, graphene is charge neutral (Fig. 1b). From the observed full-width at half-maximum of $R(V_G)$, we estimate the density inhomogeneity of the suspended graphene to be less than $1 \times 10^{10}$ cm$^{-2}$ (refs 19, 21).

We first confirm the high quality of the suspended graphene devices by studying them at low magnetic fields. Remarkably, even at $B = 0.3$ T we observe a developing plateau in $G$ at $2e^2/h$ associated with the IQHE at $v = 2$. At $B = 1$ T, fully developed plateaus are evident at $v = \pm 2, \pm 6, \pm 10$ (Fig. 1c). It is convenient to visualize the IQHE by means of a Landau fan diagram, where the derivative $|dG/dn(B,n)|$ is plotted as a function of $B$ and $n$ (Fig. 1c). At a quantum Hall plateau corresponding to filling factor $v$, the conductance remains constant $G(V_G) = v e^2/h$, and the electron density is given by $n(B) = Bv/\varphi_0$, where $\varphi_0 = h/e$ is the magnetic flux quantum. Therefore, quantum Hall plateaus $dG/dn = 0$ appear as stripes fanning out from point $B = 0$, $V_G = V_D$ and with slope $dn/dB = v/\varphi_0$. We observe such stripes down to $B < 0.1$ T, indicating that the quantum Hall effect survives to very low fields. Assuming $\mu B \gg 1$ for the IQHE to exist[19], this observation implies a lower bound for the mobility of $\sim 100,000$ cm$^2$ V$^{-1}$ s$^{-1}$.

In such ultraclean graphene samples, the extreme quantum limit can be reached at relatively low magnetic field. In the intermediate field regime $B < 10$ T (Fig. 2a), we observe an integer quantum Hall plateau corresponding to $v = 1$ at a field as low as $B = 2$ T, in addition to the well-resolved $v = 2$ integer quantum Hall plateau already mentioned. Similarly, a $G = 0$ plateau appears around the region of zero density ($V_G = V_D$) for $B > 5$ T, indicating the onset of an insulating state at low density. In the multi-terminal Hall bar samples, this insulating behaviour is considered a signature of the $v = 0$ IQHE, as it accompanies a plateau of Hall conductivity $\sigma_{xy} = 0$. In samples on a substrate[5,6], such $v = 0, \pm 1$ integer quantum Hall plateaus could only be observed for $B > 20$ T. These plateaus are believed to arise

from a magnetic field induced spontaneous symmetry breaking mediated by e–e interactions[16]. This suggests that e–e correlations in ultraclean suspended graphene can be sustained at much lower fields, compared to graphene on substrates. The strength of e–e interactions in a filled Landau level can be estimated to be proportional to $e^2/l_b$ (ref. 18), where $l_b = (\hbar/eB)^{1/2}$ is the magnetic length, and thus increases with $B$. We therefore expect stronger interaction effects to occur at higher magnetic fields.

Indeed, two notable features emerge in the high magnetic field regime, $B > 10$ T: a peak-like structure and a plateau-like structure, marked by 'D' and 'A', respectively (Fig. 2). While the position of the peak-like feature D remains unchanged, feature A moves to higher densities with increasing $B$, developing into a well-defined plateau. As the density $n$ corresponding to D does not change with $B$, this feature is probably due to $B$-independent rapid threshold fluctuations, associated with hopping or resonant tunnelling through localized states[23]. On the other hand, the electron density corresponding to the plateau-like structure A changes with increasing $B$, suggesting that A is related to a quantum Hall state with $v > 0$ (ref. 24).

A consistent picture emerges when we plot $G$ as a function of filling factor estimated as $v = n\varphi_0/B$ (Fig. 2b, inset). The traces $G(v)$ acquired at different magnetic fields collapse onto a single universal curve around the feature A. The associated plateau has a value of $G = 0.32 \pm 0.02$ $e^2/h$ and is positioned around $v = 0.30 \pm 0.02$. In contrast, the region around D does not collapse onto a universal curve, and is therefore of different origin. We assign the observed plateau A to the hitherto unobserved[7–14] fractional quantum Hall state at $v = 1/3$. We note that at our highest field, $B = 14$ T, we observe the $v = 1/3$ plateau at $n \approx 10^{11}$ cm$^{-2}$, more than an order of magnitude higher than the estimated density inhomogeneity in our device. This establishes that our observation of fractional quantization is not caused by the addition of quantum resistance observed in graphene p–n–p junctions[25,26].

Figure 3a shows our analysis of the entire data set in terms of a Landau fan diagram over the whole experimentally accessible range of $0 < B < 14$ T and $|n| < 3.5 \times 10^{11}$ cm$^{-2}$. While features stemming from localized states change between thermal cycles, the familiar integer quantum Hall states at $v = \pm 1, \pm 2, \pm 6$, appear consistently as the blue fan stripes with slope $dB/dn = \varphi_0/v$. In addition to the $v = 1/3$ fractional quantum Hall state A, we find features fanning out with distinctly different slopes corresponding to $v < 1$. These features, marked 'B' and 'C', appear only at low density and at high fields, yielding kink-like structures in $G(n)$ (Fig. 3b). We extract $v = 0.46 \pm 0.02$, and $v = 0.68 \pm 0.05$ for the states B and C, respectively, from the slopes of the corresponding lines in Fig. 3a. As before, $G(n)$ collapses onto a universal curve around B and C at different fields, allowing the extraction of $G = 0.54 \pm 0.02$ $e^2/h$ and $0.94 \pm 0.02$ $e^2/h$, where the value of conductance is taken at the centre of each feature.

These features, observed in multiple samples both on electron and hole sides (see Supplementary Information for more details), suggest additional fractional quantum Hall states, despite being less developed than the plateau at $v = 1/3$ (feature A). As the conductance values of each features are 20–40% larger than the expected $v e^2/h$ with $v = 0.46$ (feature B) and 0.68 (feature C), it is difficult to assign them to particular fractional quantum Hall states with certainty. We note however, that features B and C occur near $v = 1/2$ and $v = 2/3$. It is likely that the feature C is related to the $v = 2/3$ fractional quantum Hall state, which is expected to be one of the strongest such states in graphene[8]. In that case, the deviation of the conductance from the expected value can be ascribed to a sample-specific effect of the two-terminal measurement, namely, the mixing of longitudinal and transverse conductivity, which can yield $G$ values higher than those of traditional four-probe measurements[27]. The origin of feature B is even more unclear. Judging from FQHEs in traditional 2DESs, the next state to expect is $v = 2/5$. Yet its position around $v = 1/2$ rather suggests it might be the two-probe signature of a broad minimum in
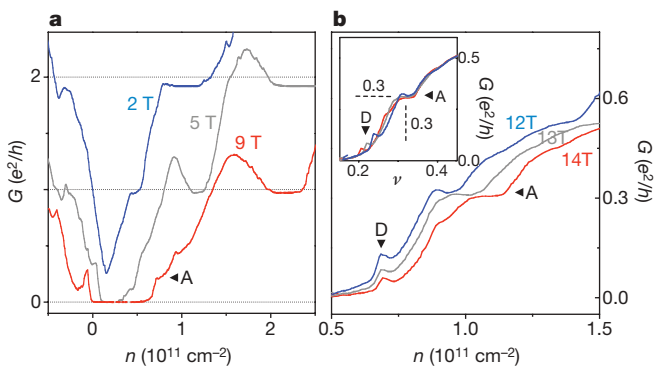


**Figure 2 | Magnetotransport at high magnetic fields. a**, $G(n)$ at $B = 2, 5, 9$ T. Quantum Hall state $v = 1$ develops at $B = 2$ T, features with $v < 1$ form at 9 T. **b**, $G(n)$ for $B = 12, 13, 14$ T. The data are acquired at $T = 6$ K to suppress resistance fluctuations associated with quantum interference. A plateau 'A' with $G \approx 0.3$ $e^2/h$ and at a density dependent on $B$ emerges when $B > 11$ T. Another feature, 'B', appears at constant $n$ at different fields. Inset, $G$ is replotted as a function of filling factor $v = n\varphi_0/B$ for $B = 12, 13, 14$ T. At different fields, the $G(v)$ curves around A collapse onto a single universal trace, allowing identification of feature A as the fractional quantum Hall state $v = 1/3$.

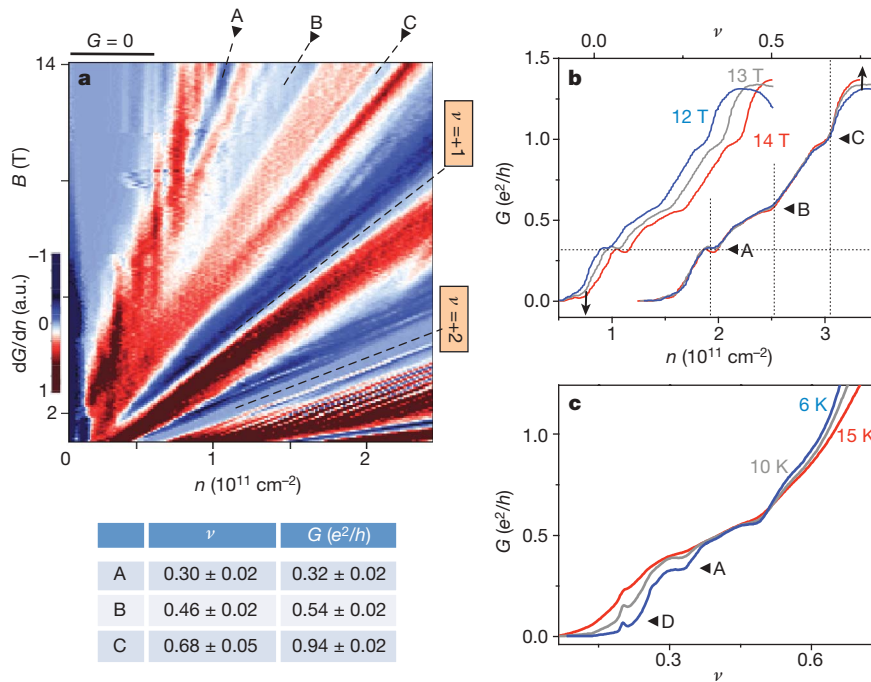| | $\nu$ | $G\ (e^2/h)$ |
|---|---|---|
| A | $0.30 \pm 0.02$ | $0.32 \pm 0.02$ |
| B | $0.46 \pm 0.02$ | $0.54 \pm 0.02$ |
| C | $0.68 \pm 0.05$ | $0.94 \pm 0.02$ |

**Figure 3 | Identifying additional fractional quantum Hall states. a**, Landau fan diagram $|dG/dn(B,n)|$ at $T = 6\,\mathrm{K}$. Dark blue colour corresponds to $G(n) = \nu e^2/h = \mathrm{const.}$ At high fields, features 'B' and 'C', with $\nu = 0.46, 0.68$, respectively, emerge in addition to feature A ($\nu = 1/3$). Insulating $G = 0$ state is evident around $n = 0$ at $B > 5\,\mathrm{T}$. The data are acquired from the same sample as Fig. 2, but in a different thermal cycle, where disorder related feature 'D' is absent. For clarity, only the electron part of the data set $n > 0$ is shown, the data for the hole side is shown in Supplementary Information. **b**, $G$ as a function of both $n$ (bottom axis) and filling factor $\nu$ (top axis) at $B = 12, 13, 14\,\mathrm{T}$. Features A, B and C in curves at different fields collapse onto a single universal curve $G(\nu)$. **c**, $G(\nu)$ at different temperatures $T = 6$, 10, 15 K. Feature A ($\nu = 1/3$) survives up to $T = 10\,\mathrm{K}$. Inset, filling factors $\nu$ and conductance $G$ values corresponding to features A, B and C.

$\rho_{xx}$ often observed in traditional 2DESs in the vicinity of $\nu = 1/2$, and not a fractional quantum Hall state[28].

The features A, B, and C persist as the temperature is increased. While most other features associated with localized states disappear at $T \approx 6\,\mathrm{K}$, the plateau corresponding to $\nu = 1/3$ survives up to $T \approx 10\,\mathrm{K}$ (Fig. 3c). This energy scale is higher than the energy scale of the FQHE observed in typical semiconductor heterojunctions with similar mobility, where the 1/3 plateau disappears for $T > 1\,\mathrm{K}$ at $B = 15\,\mathrm{T}$ (ref. 1). This remarkable resilience of the FQHE in graphene can be ascribed to the enhanced e–e interaction in suspended graphene due to a reduced dielectric screening. The dielectric constant $\varepsilon$ enters the characteristic energy scale of e–e interaction as $E_{e\text{-}e} = e^2/$ $\varepsilon l_b$. In suspended graphene $\varepsilon \approx 1$, as compared to $\varepsilon = 13.6$ in AlGaAs/ GaAs (ref. 18), leading to much stronger e–e interactions in graphene, and correspondingly, to larger energy gaps and more robust FQHEs. In fact, the theoretical expectation for the excitation energy of $\nu = 1/3$ FQHE in graphene is $0.017 e^2/\varepsilon l_b \approx 20\,\mathrm{K}$ (ref. 8), which agrees with our observations.

Despite the lack of experimental observation so far, many theorists have addressed the possibility of the FQHE in graphene. When only SU(2) symmetry is assumed[8,10] (which is appropriate for the case of high Zeeman energy), it was found that the FQHE in graphene maps on to the well-studied problem of the FQHE in GaAs 2DESs[2]. Indeed, the fractional quantum Hall states observed in the present work,
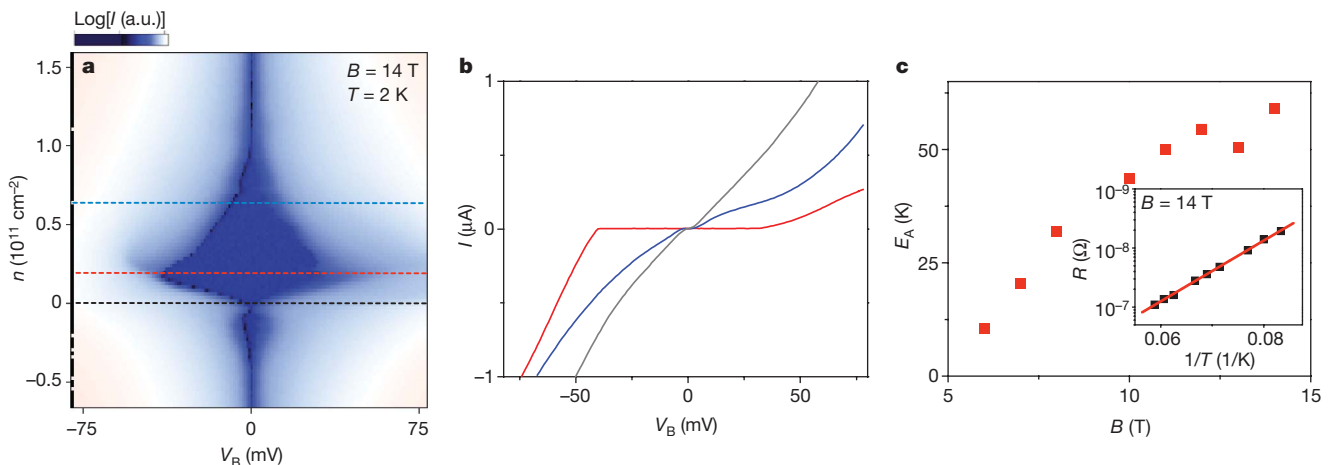


**Figure 4 | The insulating state in graphene near zero density. a**, Colour rendition of logarithm of the absolute value of current $I$ across the device as a function of bias voltage $V_B$ at $B = 14\,\mathrm{T}$ and $T = 2\,\mathrm{K}$. The insulating state $I = 0$ is evident as a dark blue region near $V_B = 0$. **b**, $I(V_B)$ at $n = 0$ (red), $0.17 \times 10^{11}\,\mathrm{cm}^{-2}$ (blue) and $1 \times 10^{11}\,\mathrm{cm}^{-2}$ (black). **c**, Activation energy $E_A$ at $n = 0.17 \times 10^{11}\,\mathrm{cm}^{-2}$ as a function of applied magnetic field $B$. Inset, logarithm of the resistance $R$ versus inverse temperature $1/T$ at $B = 14\,\mathrm{T}$ and $n = 0$. The linear fit yields $R \propto \exp(E_A/2kT)$ with $E_A \approx 60\,\mathrm{K}$.

$v = 1/3$ and possibly $v = 2/3$, are in fact the strongest states in the GaAs FQHE. When, however, the full SU(4) symmetry is considered, novel fractional quantum Hall states without GaAs analogues were predicted[7,13,14]. We expect that better quality specimens will allow exploration of these currently inaccessible exotic possibilities.

Finally, we consider the insulating state that appears in the low density regime at $|v| < 1/3$. The device is fully insulating ($R > 10\,\mathrm{G}\Omega$) at magnetic fields $B > 5\,\mathrm{T}$ and at filling factors $v < 0.15$. In the Landau fan diagram (Fig. 3a) this insulating state is apparent as a region of zero conductance $\mathrm{d}G/\mathrm{d}n = 0$ near the charge neutrality point. To elucidate the nature of this insulating state, we examine the temperature dependence of the device's resistance measured with a small bias voltage ($V_b < 0.1\,\mathrm{mV}$). Whereas at the highest attainable field ($B = 14\,\mathrm{T}$) and low temperatures ($T < 10\,\mathrm{K}$) the device is fully insulating, at higher temperatures we observe a clear Arrhenius behaviour, $R(n = 0.17 \times 10^{11}\,\mathrm{cm}^{-2}) \propto \exp(E_A/2kT)$, with an activation energy $E_A \approx 60\,\mathrm{K}$ (Fig. 4c inset). Similar analyses at different magnetic fields reveal that $E_A$ rapidly decreases with decreasing $B$ (Fig. 4c).

Such a highly resistive state has been previously observed in more disordered graphene on substrates[17]. Although its nature is hotly debated[16], the consensus is that this state stems from the symmetry breaking of the zeroth Landau level by e–e interactions. The activated behaviour clearly observed in our ultraclean suspended graphene suggests the presence of a gap in the density of extended states near the Dirac point. Such a gap opening can also be probed by measuring the current $I$ as a function of the source–drain bias voltage, $V_B$. Figure 4a shows $\log|I|$ as a function of both $V_B$ and $V_G$. A dark region in the central portion of this plot near $V_G = V_D$ indicates a region with near-zero conductance. Several representative $I$–$V_B$ curves at different densities (Fig. 4b) exhibit distinctly nonlinear behaviour. In particular, near the charge neutrality point $V_G = V_D$, $I$ remains close to zero over a wide bias range $|V_b| < 50\,\mathrm{mV}$. This is much larger than the energy scale corresponding to the activation energy, $\sim 4\,\mathrm{meV}$, suggesting that at large bias the charge transport occurs across weakly connected insulating regions.

1. Tsui, D. C., Stormer, H. L. & Gossard, A. C. Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.* **48**, 1559–1562 (1982).
2. Jain, J. *Composite Fermions* (Cambridge Univ. Press, 2007).
3. Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197–200 (2005).
4. Zhang, Y., Tan, Y.-W., Stormer, H. L. & Kim, P. Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature* **438**, 201–204 (2005).
5. Jiang, Z., Zhang, Y., Stormer, H. L. & Kim, P. Quantum Hall states near the charge-neutral Dirac point in graphene. *Phys. Rev. Lett.* **99**, 106802 (2007).
6. Zhang, Y. *et al.* Landau-level splitting in graphene in high magnetic fields. *Phys. Rev. Lett.* **96**, 136806 (2006).
7. Goerbig, M. O. & Regnault, N. Analysis of a SU(4) generalization of Halperin's wave function as an approach towards a SU(4) fractional quantum Hall effect in graphene sheets. *Phys. Rev. B* **75**, 241405 (2007).
8. Toke, C., Lammert, P. E., Crespi, V. H. & Jain, J. K. Fractional quantum Hall effect in graphene. *Phys. Rev. B* **74**, 235417 (2006).
9. Peres, N. M. R., Guinea, F. & Neto, A. H. C. Electronic properties of disordered two-dimensional carbon. *Phys. Rev. B* **73**, 125411 (2006).
10. Apalkov, V. M. & Chakraborty, T. Fractional quantum Hall states of Dirac electrons in graphene. *Phys. Rev. Lett.* **97**, 126801 (2006).
11. Yang, K., Sarma, S. D. & MacDonald, A. H. Collective modes and skyrmion excitations in graphene SU(4) quantum Hall ferromagnets. *Phys. Rev. B* **74**, 075423 (2006).
12. Khveshchenko, D. V. Composite Dirac fermions in graphene. *Phys. Rev. B* **75**, 153405 (2007).
13. Shibata, N. & Nomura, K. Coupled charge and valley excitations in graphene quantum Hall ferromagnets. *Phys. Rev. B* **77**, 235426 (2008).
14. Toke, C. & Jain, J. K. SU(4) composite fermions in graphene: fractional quantum Hall states without analog in GaAs. *Phys. Rev. B* **75**, 245440 (2007).
15. Neto, A. H. C., Guinea, F., Peres, N. M. R., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109–163 (2009).
16. Yang, K. Spontaneous symmetry breaking and quantum Hall effect in graphene. *Solid State Commun.* **143**, 27–32 (2007).
17. Checkelsky, J. G., Li, L. & Ong, N. P. Divergent resistance at the Dirac point in graphene: evidence for a transition in a high magnetic field. *Phys. Rev. B* **79**, 115434 (2009).
18. Sarma, S. D. & Pinczuk, A. *Perspectives in Quantum Hall Effects: Novel Quantum Liquids in Low-Dimensional Semiconductor Structures* (Wiley, 1997).
19. Bolotin, K. I. *et al.* Ultrahigh electron mobility in suspended graphene. *Solid State Commun.* **146**, 351–355 (2008).
20. Du, X., Skachko, I., Barker, A. & Andrei, E. Y. Approaching ballistic transport in suspended graphene. *Nature Nanotechnol.* **3**, 491–495 (2008).
21. Bolotin, K. I., Sikes, K. J., Hone, J., Stormer, H. L. & Kim, P. Temperature-dependent transport in suspended graphene. *Phys. Rev. Lett.* **101**, 096802 (2008).
22. Nomura, K. & MacDonald, A. H. Quantum Hall ferromagnetism in graphene. *Phys. Rev. Lett.* **96**, 256602 (2006).
23. Cobden, D. H., Barnes, C. H. W. & Ford, C. J. B. Fluctuations and evidence for charging in the quantum Hall effect. *Phys. Rev. Lett.* **82**, 4695–4698 (1999).
24. Martin, J. *et al.* The nature of localization in graphene under quantum Hall conditions. *Nature Phys.* **5**, 669–674 (2009); published online 26 July 2009.
25. Ozyilmaz, B. *et al.* Electronic transport and quantum Hall effect in bipolar graphene p-n-p junctions. *Phys. Rev. Lett.* **99**, 166804 (2007).
26. Williams, J. R., DiCarlo, L. & Marcus, C. M. Quantum Hall effect in a gate-controlled p-n junction of graphene. *Science* **317**, 638–641 (2007).
27. Abanin, D. A. & Levitov, L. S. Conformal invariance and shape-dependent conductance of graphene samples. *Phys. Rev. B* **78**, 035416 (2008).
28. Willett, R. L., West, K. W. & Pfeiffer, L. N. Transition in the correlated 2D electron system induced by a periodic density modulation. *Phys. Rev. Lett.* **78**, 4478–4481 (1997).

**Author Contributions** K.I.B. and F.G. performed the experiments and analysed the data. M.D.S. assisted with fabrication. H.L.S. and P.K. conceived the project. All authors contributed to writing the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.K. (pk2015@columbia.edu).

# LETTERS

# Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy

Chong Fang[1], Renee R. Frontiera[1], Rosalie Tran[1] & Richard A. Mathies[1]

Tracing the transient atomic motions that lie at the heart of chemical reactions requires high-resolution multidimensional structural information on the timescale of molecular vibrations, which commonly range from 10 fs to 1 ps. For simple chemical systems, it has been possible to map out in considerable detail the reactive potential-energy surfaces describing atomic motions and resultant reaction dynamics[1], but such studies remain challenging for complex chemical and biological transformations[2]. A case in point is the green fluorescent protein (GFP)[3–5] from the jellyfish *Aequorea victoria*, which is a widely used gene expression marker owing to its efficient bioluminescence. This feature is known to arise from excited-state proton transfer (ESPT)[6–8], yet the atomistic details of the process are still not fully understood. Here we show that femtosecond stimulated Raman spectroscopy[9,10] provides sufficiently detailed and time-resolved vibrational spectra of the electronically excited chromophore of GFP to reveal skeletal motions involved in the proton transfer that produces the fluorescent form of the protein. In particular, we observe that the frequencies and intensities of two marker bands, the C–O and C=N stretching modes at opposite ends of the conjugated chromophore, oscillate out of phase with a period of 280 fs; we attribute these oscillations to impulsively excited low-frequency phenoxyl-ring motions, which optimize the geometry of the chromophore for ESPT. Our findings illustrate that femtosecond simulated Raman spectroscopy is a powerful approach to revealing the real-time nuclear dynamics that make up a multidimensional polyatomic reaction coordinate.

The wild-type GFP chromophore embedded in the protein β-barrel structure is formed by the autocatalytic cyclization of the Ser 65-Tyr 66-Gly 67 tripeptide[11,12]. The hydrogen-bonding network (Fig. 1a) linking the two ends of the chromophore includes the hydroxyl of Tyr 66, a water molecule (Wat 22), the side-chain hydroxyl of Ser 205, the carboxylate of Glu 222 and the side-chain hydroxyl of Ser 65. Following excitation with light of wavelength 397 nm, the excited neutral chromophore (A*) converts to the unrelaxed anionic form (I*) by transferring the phenolic proton from Tyr 66 to Glu 222 on the picosecond timescale[6,7]. The chromophore in GFP is thus a classic example of a Kasha-type ESPT reaction[13]. Whereas the model chromophore in solution or denatured protein exhibits extremely low fluorescence at room temperature[14,15], wild-type GFP emits very strong green fluorescence once the proton transfer occurs. However, the nuclear motions of the chromophore that lead to ESPT in GFP are not yet known.

Photochemical-reaction dynamics occurring on the femtosecond timescale in biological systems such as rhodopsin[16], bacteriorhodopsin[17] and phytochrome[18] have been successfully studied with femtosecond stimulated Raman spectroscopy (FSRS). This vibrational structural technique[9] uses a femtosecond actinic pump pulse to initiate photochemistry, and a picosecond Raman pump pulse together with a 20-fs Raman probe pulse to stimulate broadband Raman signals within a possible vibrational range of 50–3,500 cm$^{-1}$. The time resolution of FSRS is determined by

the cross-correlation between the femtosecond actinic pump and Raman probe pulses, and can be 25 fs or less. The energy resolution is determined by the Raman pump bandwidth and the vibrational-coherence free-induction decay time, and can be as narrow as 5 cm$^{-1}$.

Figure 1b shows vibrational spectra recorded with FSRS of the wild-type GFP neutral chromophore in $H_2O$ buffer and excited at 397 nm. The A* spectrum appears with a 120-fs rise time and decays to I* on the picosecond timescale. It has intense peaks assigned to the C–H rocking (1,100–1,220 cm$^{-1}$) and C–O stretching (1,280 cm$^{-1}$) modes on the phenol ring and the C–N stretching (1,400–1,480 cm$^{-1}$) and C=N stretching (1,585 cm$^{-1}$) modes on the imidazolinone ring. The C=N stretching mode exhibits a redshift of 20 cm$^{-1}$ as time passes (Supplementary Fig. 1; a red-shifted C=O stretching mode may also contribute to the band). This redshift, with a time constant of $5.1 \pm 0.2$ ps, is indicative of a structural transition assigned to ESPT and involving multiple proton motions through the proton wire[6,7] (Fig. 1a). FSRS data in $D_2O$ yield a similar redshift in the C=N stretching mode, but with a much greater time constant, $22 \pm 2$ ps. The time constants give a kinetic isotope effect (KIE) of $4.3 \pm 0.6$, which is consistent with the KIE of 5 reported from A* emission decay[7] but smaller than the KIE of 6 derived from the appearance of the I* state[19].

The plot of the intensity of the phenol C–H rocking mode as a function of time (Fig. 1c) shows two distinct relaxation phases. The subpicosecond components of the A* decay (Supplementary Discussion) in the two solvents are comparable (700 fs in $H_2O$ and 600 fs in $D_2O$), but the long-time decay constant is significantly greater in $D_2O$. The similar initial decays show that solvent deuteration does not modify the Franck–Condon dynamics on the A* potential-energy surface (PES), suggesting that proton motion (essential for product formation) is not a dominant part of the 600–700-fs decay process. Nevertheless, C–O vibrational features characteristic of the I* product emerge within 2 ps in $H_2O$ and within 5 ps in $D_2O$ (Fig. 1c), indicating a short dwell between the fast initial decay process and the first product appearance. The ultrafast structural rearrangements allowing ESPT and product formation are triggered by electronic changes associated with the excitation ($\pi \rightarrow \pi^*$) from the highest occupied molecular orbital to the lowest unoccupied molecular orbital, which, according to calculations[20] illustrated[21] in Fig. 1d, involve limited electron migration from the phenol to the imidazolinone ring with the electron density change localized on the phenol ring and the bridging bonds between the two rings ($|\Delta\mu| \approx 2.5$ D, where $\mu$ is the static electric dipole moment; see Supplementary Fig. 2 for spectral changes).

Figure 2a shows a series of FSRS spectra of the chromophore in $D_2O$ buffer similar to those in Fig. 1b, but with the individual spectra separated by smaller time steps to allow transient structural changes to be explored more precisely. Remarkably, the vibrational-mode frequencies are not constant and their intensities do not monotonically decay with time; instead, a number of modes have periodic modulations or

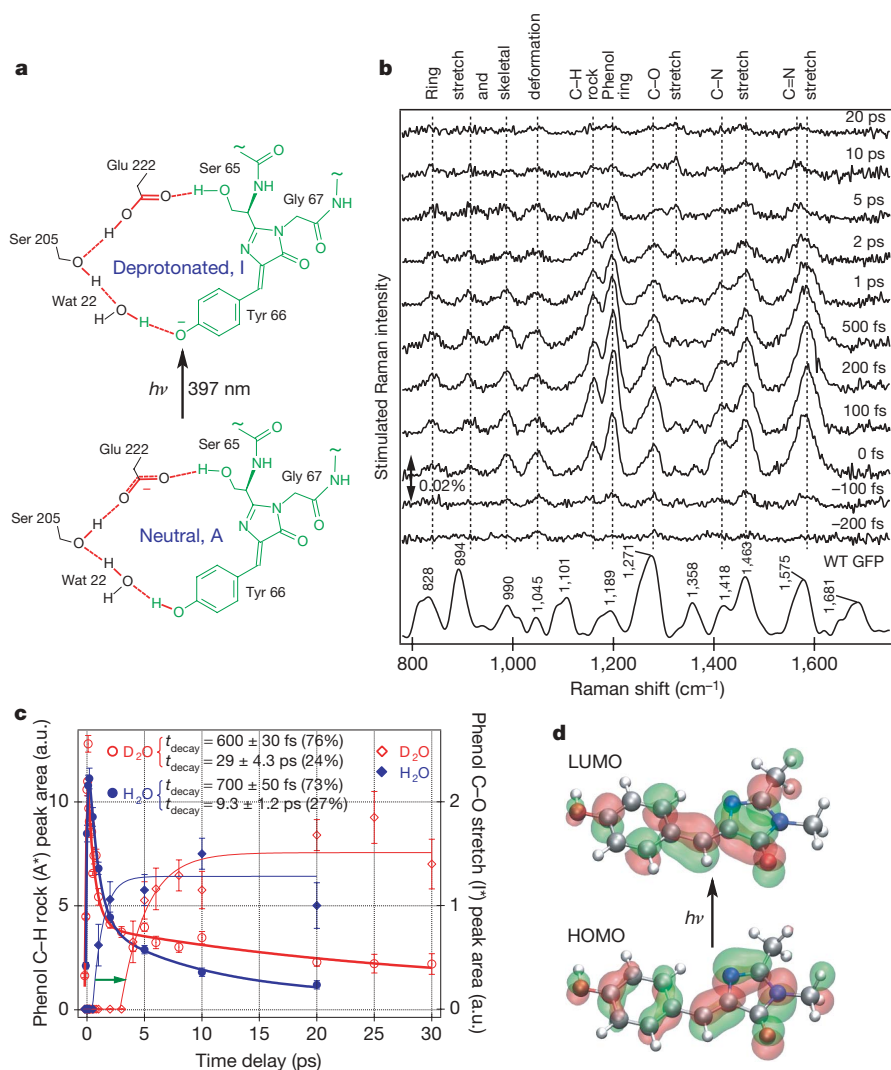[1]Department of Chemistry, University of California, Berkeley, California 94720, USA.

**Figure 1 | Ultrafast structural response of the photoexcited wild-type GFP chromophore.**
**a**, The hydrogen-bonding network linking the two ends of the chromophore in the protein pocket before (neutral, A) and after (deprotonated, I) proton transfer. The tildes indicate the omission of the protein residues not directly involved in the ESPT chain. **b**, Femtosecond stimulated Raman spectra of wild-type (WT) GFP in $H_2O$ from −200 fs to 20 ps after 397-nm excitation. The spectra are vertically offset and have had the ground-state spectrum (shown at bottom; stimulated using a 794-nm Raman pump pulse) and a baseline removed. **c**, Plot of the intensity of the A* phenol C–H rocking mode (circles) as a function of time, revealing a two-phase decay process, with the proton-transfer component significantly slowed by solvent deuteration; $t_{decay}$ is the decay time constant from data fitting, and the percentages in parentheses indicate the relative weights of individual fitted components. The intensity of the I* phenol C–O stretching mode (diamonds) is displayed on the right-hand axis, for protonated and deuterated solvents. Error bars, s.d. ($n = 3$). a.u., arbitrary units. **d**, Highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals of the neutral model chromophore *in vacuo* as calculated using density functional theory (DFT) at the B3LYP/6-31G* level[20]. Carbon, oxygen, nitrogen and hydrogen atoms are shown as grey, red, blue and white spheres, respectively. Red and green shading distinguishes the positive and negative phases of the molecular orbitals, respectively, displayed with an isovalue of $0.03e\,\mathring{A}^{-3}$.

quantum beats on the timescale of a few hundred femtoseconds (see Supplementary Fig. 3 for a detailed analysis). For example, the phenol C–H rocking modes at 1,145 and 1,190 $cm^{-1}$ and the C–N and C=N stretching modes at 1,405, 1,450 and 1,565 $cm^{-1}$, which are located at opposite ends of the two conjugated rings of the chromophore, exhibit out-of-phase intensity oscillations, as highlighted by the horizontal dashed lines in Fig. 2b.

Figure 3 presents additional plots of the frequencies and intensities of the phenol and imidazolinone ring modes as functions of time. Figure 3a shows that two of the modes on the phenol ring (the C–O stretching mode and the phenol-ring deformation mode) display in-phase frequency oscillations that have a period of ~280 fs and are out of phase with the modes on the imidazolinone ring (the C=N stretching mode and the imidazolinone-ring deformation mode). The phenolic (C)O–D rocking mode at ~1,100 $cm^{-1}$ also has large-scale oscillations with a period of ~280 fs, but they are in antiphase with the C–O stretching mode at 1,262 $cm^{-1}$ on the phenol ring because the shift in electron density from the phenolic hydroxyl to the phenol ring causes a concomitant bond-order increase in C–O and decrease in O–D. We note that the phenol-ring mode frequencies increase after time zero, whereas the imidazolinone-ring deformation mode displays the opposite trend. The overlay of the intensity oscillations of the C–O and C–N stretching modes (Fig. 3b) shows that they begin to oscillate in opposite directions after time zero, confirming that the photoexcitation-induced π-electron conjugation changes affect the two ring modes in opposite ways. As the chromophore modifies its conformation in the local environment mediated by the external proton-relay wire, the chemical bonds periodically vary their strengths and polarizabilities

in response to the dynamic structural changes controlling the electron distribution over the two rings. Fourier-transform analysis of these time-resolved oscillations yields a common low-frequency mode at 120 ± 20 $cm^{-1}$. Similar spectral oscillations with a period of 270 ± 50 fs (125 ± 25 $cm^{-1}$) are also observed in $H_2O$.

To understand the early events on the A* PES in more detail, actual molecular frequency changes have to be determined. In FSRS, the observed frequency modulation is smaller than the intrinsic molecular magnitude because the vibrational features are averaged over the vibrational dephasing time. The relevant 350-fs dephasing time of A* is determined by the average peak width, 45 $cm^{-1}$. On the basis of coupled-wave theory[9], we simulated the macroscopic polarization and least-squares fitted the time-resolved data (Fig. 4a), to find that the intrinsic frequency of the C–O stretching mode in $D_2O$ oscillates between ~1,244 and ~1,276 $cm^{-1}$ with a period of 280 fs. These oscillations are a direct consequence of the temporal variation of the degree of hydrogen bonding at the phenolic hydroxyl group as the GFP chromophore tautomerizes between a tyrosine and a tyrosinate-like form on the femtosecond timescale. They cover a frequency range of 32 $cm^{-1}$, which is similar in magnitude to the observed blueshift, of 38 $cm^{-1}$, that occurs in the C–O stretching mode after the phenolic deuteron has been fully transferred in $D_2O$, indicating that the observed oscillatory motion moves the system significantly along the ESPT reaction coordinate.

Low-frequency vibrational modes have been shown to be coupled to the reaction coordinate of biological processes[22,23]. In the case of GFP, DFT calculations on the neutral chromophore 4′-hydroxybenzylidene-2,3-dimethylimidazolinone predict a 110-$cm^{-1}$ two-ring out-of-plane
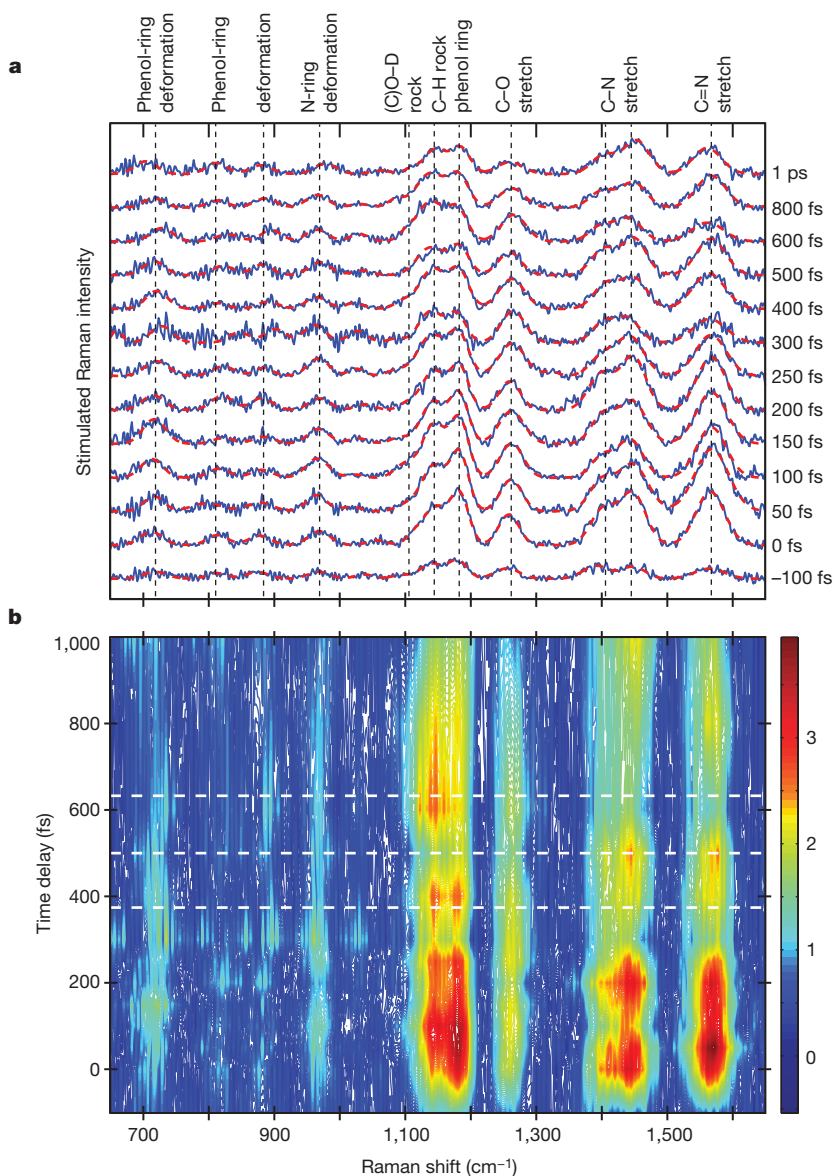
a



b



**Figure 2 | FSRS spectra of wild-type GFP in $D_2O$ from −100 fs to 1 ps following 396-nm excitation. a,** Gaussian fits are shown as dashed red lines on each delay time spectrum (blue). **b,** Two-dimensional plot of the FSRS spectra with 100 contour lines drawn from peak minimum to maximum (colour scale, a.u.). Intensity and frequency oscillations of the indicated vibrational modes are apparent. The dashed white lines serve as a visual guide for the out-of-phase intensity modulations of the C–H rocking, C–O stretching and C–N stretching, C=N stretching modes located on the opposite rings of the GFP chromophore.

wagging mode on the electronic ground state, $S_0$, the frequency of which should not change significantly upon excitation or exchange of hydrogen and deuterium in the buffer. The crystal structure of wild-type GFP[6] shows that Wat 22 is located ~1.3 Å above the plane of the two rings and to one side of the phenolic end of the chromophore. The phenolic hydroxyl is initially hydrogen-bonded to Wat 22 (Supplementary Discussion) but there is a large barrier for deprotonation in $S_0$ (Fig. 4b; structure I, grey). Photoexcitation lowers the $pK_a$ of Tyr 66 and shifts some electron density from the phenol ring to the imidazolinone ring, making the phenolic proton more acidic. The impulsively excited wagging motions cause the phenol ring to swing upward (Fig. 4b; structure II, orange), away from the originally planar structure, shifting the phenolic proton towards Wat 22 and strengthening the existing hydrogen bond. The initial increase of the phenolic C–O double-bond character and Raman peak intensity (Fig. 3) provides direct evidence for this motion. Concomitantly, the π-conjugation over the two rings shifts to resemble a more deprotonated chromophore with a lengthened C=N bond. This tautomeristic transformation is hence structurally and temporally resolved. As time passes and the phenol ring oscillates back to below the original plane (Fig. 4b; structure III, cyan), the phenolic proton shifts away from Wat 22 and the frequency of the C–O stretching mode decreases. The large-amplitude oscillation in the frequency of the phenolic (C)O–D rocking mode (Fig. 3a) shows that it also is strongly affected by this motion, which is readily explained by the

deuteron–Wat 22 distance and geometry being highly perturbed by the wagging of the phenol ring. The build-up of the spectral oscillations with time suggests that there is a significant increase in the electronic polarizability and its vibrational modulation as the wavepacket moves along the reaction coordinate towards the transition state, where the transportation of the deuteron begins. Thus, the ~120-cm$^{-1}$ skeletal mode is important in changing the system to a geometry that facilitates rearrangements of hydrogen bonds (Supplementary Discussion).

Other low-frequency modes are also probably impulsively excited and involved in the reaction coordinate, but the wagging mode is the most dominant of the experimentally observed contributions. In particular, ESPT may also occur in concert with the phenol-ring wagging at the imidazolinone ring, which is in contact with Glu 222 through a hydrogen bond. The derived frequency modulations for the C–O and C=N stretching modes show there to be an unusually large anharmonic coupling[24] between the coherent low-frequency modes and the high-frequency bond stretching modes on the multidimensional excited-state PES of GFP; it is this anharmonicity that leads to efficient proton transfer.

The functional role of the ring wagging dynamics shows that the ground-state hydrogen-bond geometry is not optimized for ESPT; if it were, the phenolic proton would be transferred immediately after excitation. The lack of a KIE during the first 1–2 ps of oscillatory A* decay supports the interpretation that the initial reaction coordinate
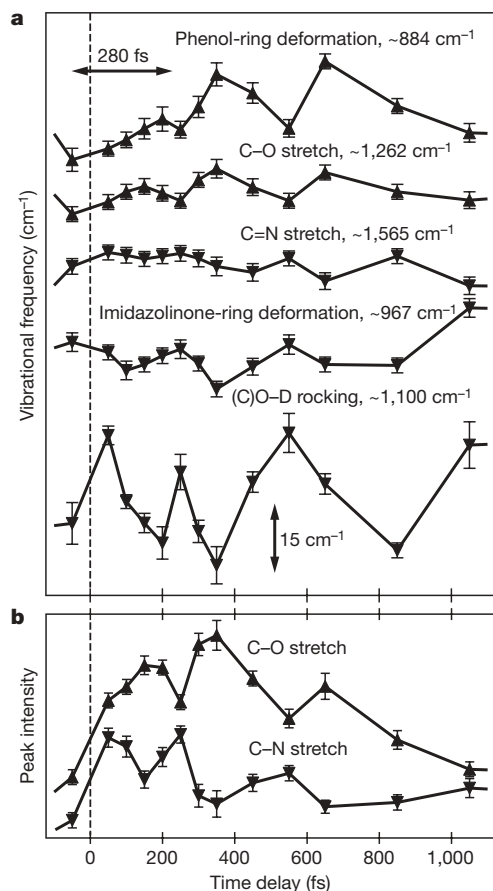
**Figure 3 | Time-dependent vibrational structural features of the excited wild-type GFP chromophore in D$_2$O. a**, The mode frequencies oscillate with a period of ~280 fs. The C=N stretching, imidazolinone-ring deformation and (C)O–D rocking modes are in phase, and this group of modes oscillates out of phase with the C–O stretching and phenol-ring deformation modes. Error bars, s.d. of the Gaussian-fitted peak centre frequency ($n = 3$). **b**, The integrated peak area oscillations of the C–O and C–N stretching modes have a period of ~280 fs and are also ~180° out of phase. Error bars, s.d. ($n = 3$).



**Figure 4 | Deconvolved marker-band frequency oscillations reveal multidimensionality of the reaction coordinate of wild-type GFP. a**, Experimental traces of the time-dependent frequencies of the C–O and C=N stretching modes in D$_2$O are in red, normalized oscillations with the broad sloping component removed are in blue, and the fits are in green. The oscillation amplitude of the intrinsic molecular frequency is ~32 cm$^{-1}$ (black), consistent with the process of making and breaking transient hydrogen bonds. The two marker-band motions, localized at opposite ends of the chromophore, oscillate out of phase. **b**, Perspective view of the phenol ring-wagging motion that modulates the chromophore vibrational frequencies. Three positions (grey (I) at 0 and 140 fs, 0°; orange (II) at 70 fs, +10°; cyan (III) at 210 fs, −10°) in the phenol-ring motion are illustrated to show how the phenol ring swings from the native structure (I) towards (structure II) and away (structure III) from the optimal ESPT geometry. Given the geometrical constraints of the protein crystal structure[6,12] and the DFT-calculated mode displacements, the chromophore wagging angle of 10° was chosen for structure II to bring the phenolic proton into optimal hydrogen-bonding geometry with Wat 22. Atomic colours are as in Fig. 1d; green background, protein pocket.

does not involve significant proton motion but is instead dominated by skeletal motions[25] that enable the chromophore to search phase space for the correct orientation and trajectory within its protein environment to transfer the proton. Once the geometry of the nearby heavy atoms is largely optimized for ESPT, proton motion becomes the rate-determining step consistent with a KIE of ~4.3. However, because the maximal KIE for proton transfer reactions in the symmetric limit[26] is 6–10, the picosecond-ESPT reaction coordinate of GFP must still involve some degree of skeletal rearrangement. This is corroborated by the appearance of a peak in the C–O stretching mode that is blueshifted by 45 cm$^{-1}$ after a dwell of 2 ps between the A* oscillatory motion and the initial I* formation in H$_2$O and by 38 cm$^{-1}$ after a dwell of 5 ps in D$_2$O (Fig. 1c); during this dwell, the heavy-atom skeleton rearranges in concert with the new hydrogen-bonding configuration.

Our observations regarding the nature of the vibrational motion involved, the magnitude of the bonding changes and the temporal correlation of wagging and product appearance support a causative connection between wagging and proton transfer. The skeletal wagging mode thus has a large projection along the ESPT reaction coordinate in wild-type GFP and makes an essential contribution to the overall reaction. Although it is true that the proton transfer occurs on a timescale of about 5–20 ps and the impulsive wagging motion is unlikely to persist for more than a few picoseconds, the aforementioned functional role does not require the lifetime of the wagging

oscillation to persist until proton transfer is complete. Rather, in the first picoseconds after excitation, the net evolution along the wagging coordinate facilitates a shift in the electron density and equilibrium nuclear positions in the direction of the tautomerism, setting up a more favourable ESPT environment. We note that recent *ab initio* molecular dynamics simulations suggesting that proton transfer in methyl salicylate is regulated by molecular skeletal rearrangement[25] are consistent with this picture.

The protein environment of GFP has an essential role in orchestrating ESPT. The initial chromophore wagging is triggered by the photoexcitation-induced alleviation of π-conjugation across the bridging bonds between the two rings and the steric interactions

exerted by the surrounding protein residues that do not favour a planar chromophore[27]. In solution, the model chromophore undergoes rapid *cis–trans* isomerization, which effectively releases the excitation energy through radiationless pathways[14,15]. However, the strong geometric constraints and the well-defined hydrogen-bonding system in GFP limit chromophore motion[27]. This suggests that the protein pocket inhibits free rotation around the chromophore bridging bonds[15] and thereby blocks isomerization-based internal conversion channels; allows limited chromophore motion on photoexcitation by providing restricted space for the two-ring intramolecular wagging motion away from planarity; and uses extensive hydrogen bonds and other non-covalent interactions to maintain the chromophore near the hydrogen-bonding geometry that enables the system to undergo other, slower, changes involving high-reduced-mass motions. As a result, once the molecule evolves down the A* PES along a wagging-mediated reaction coordinate, it encounters a higher isomerization barrier but a lower ESPT barrier, at which point the proton can be efficiently transferred through the hydrogen-bonding wire to form the long-lived and highly fluorescent I* state.

The kinetic structural observations presented here indicate that the early-time ring-wagging motion of the GFP chromophore is a principal component of the ESPT reaction coordinate, and we anticipate that similar low-frequency skeletal displacements are probably ubiquitous in the reaction coordinates of many photochemical and biological reactions. Encouraged by this first experimental view of the multidimensional reaction coordinate in GFP, we note that it is technically feasible to use a photoexcitation pulse of less than 10 fs (ref. 28) and a Raman probe pulse of less than 20 fs in FSRS to excite impulsively almost all the Raman-active vibrational modes seen here and record time-resolved spectra with 10–20-fs resolution. Such experiments should reveal a rich matrix of anharmonicities as a function of time, providing an unprecedented experimental view of a multidimensional PES landscape along the chemical reaction coordinate of polyatomic molecules.

## METHODS SUMMARY

The wild-type GFP sample was expressed from plasmids containing the His-tagged wild-type GFP gene and prepared in both $H_2O$ and $D_2O$ buffer solutions at ~pH 8 following standard methods. We used the output of a 1-kHz Ti:sapphire regenerative amplifier to generate the 80-fs, 400-nJ actinic pump pulse at a wavelength of 397 nm (full-width at half-maximum, 140 cm$^{-1}$), and obtained the 20-fs, 7-nJ Raman probe pulse at 840–960 nm by continuum generation and prism compression. The cross-correlation time resolution was 130 fs. The 3-ps, 2-μJ Raman pump pulse at 794 nm was produced by spectrally filtering ~90% of the amplifier output. At the focal point of the three pulses, 600 μl of $A_{396\,nm} = 1$ sample solution was flowed through a 1-mm cell at room temperature of 22 °C. We dispersed the probe pulse carrying the stimulated Raman gain features in a spectrograph and recorded it using a dual-photodiode array, from which the time-resolved excited-state spectra were generated, processed and analysed.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Zewail, A. H. *Femtochemistry: Ultrafast Dynamics of the Chemical Bond* 217–326 (World Scientific, 1994).
2. Levine, B. G. & Martínez, T. J. Isomerization through conical intersections. *Annu. Rev. Phys. Chem.* **58**, 613–634 (2007).
3. Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea. J. Cell. Comp. Physiol.* **59**, 223–239 (1962).
4. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
5. Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544 (1998).
6. Brejc, K. *et al.* Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc. Natl Acad. Sci. USA* **94**, 2306–2311 (1997).
7. Chattoraj, M., King, B. A., Bublitz, G. U. & Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl Acad. Sci. USA* **93**, 8362–8367 (1996).
8. Stoner-Ma, D. *et al.* Proton relay reaction in green fluorescent protein (GFP): polarization-resolved ultrafast vibrational spectroscopy of isotopically edited GFP. *J. Phys. Chem. B* **110**, 22009–22018 (2006).
9. Kukura, P., McCamant, D. W. & Mathies, R. A. Femtosecond stimulated Raman spectroscopy. *Annu. Rev. Phys. Chem.* **58**, 461–488 (2007).
10. McCamant, D. W., Kukura, P., Yoon, S. & Mathies, R. A. Femtosecond broadband stimulated Raman spectroscopy: apparatus and methods. *Rev. Sci. Instrum.* **75**, 4971–4980 (2004).
11. Ormo, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).
12. Yang, F., Moss, L. G. & Phillips, G. N. Jr. The molecular structure of green fluorescent protein. *Nature Biotechnol.* **14**, 1246–1251 (1996).
13. Kasha, M. Proton-transfer spectroscopy. Perturbation of the tautomerization potential. *J. Chem. Soc. Faraday Trans. II* **82**, 2379–2392 (1986).
14. Mandal, D., Tahara, T. & Meech, S. R. Excited-state dynamics in the green fluorescent protein chromophore. *J. Phys. Chem. B* **108**, 1102–1108 (2004).
15. Weber, W., Helms, V., McCammon, J. A. & Langhoff, P. W. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proc. Natl Acad. Sci. USA* **96**, 6177–6182 (1999).
16. Kukura, P., McCamant, D. W., Yoon, S., Wandschneider, D. B. & Mathies, R. A. Structural observation of the primary isomerization in vision with femtosecond-stimulated Raman. *Science* **310**, 1006–1009 (2005).
17. McCamant, D. W., Kukura, P. & Mathies, R. A. Femtosecond stimulated Raman study of excited-state evolution in bacteriorhodopsin. *J. Phys. Chem. B* **109**, 10449–10457 (2005).
18. Dasgupta, J., Frontiera, R. R., Taylor, K. C., Lagarias, J. C. & Mathies, R. A. Ultrafast excited-state isomerization in phytochrome revealed by femtosecond stimulated Raman spectroscopy. *Proc. Natl Acad. Sci. USA* **106**, 1784–1789 (2009).
19. Kennis, J. T. M. *et al.* Uncovering the hidden ground state of green fluorescent protein. *Proc. Natl Acad. Sci. USA* **101**, 17988–17993 (2004).
20. Gaussian. 03 (Gaussian Inc., Wallingford, Connecticut, 2004).
21. Humphrey, W., Dalke, A. & Schulten, K. VMD – visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
22. Zhu, L., Sage, J. T. & Champion, P. M. Observation of coherent reaction dynamics in heme proteins. *Science* **266**, 629–632 (1994).
23. Wang, Q., Schoenlein, R. W., Peteanu, L. A., Mathies, R. A. & Shank, C. V. Vibrationally coherent photochemistry in the femtosecond primary event of vision. *Science* **266**, 422–424 (1994).
24. Kukura, P., Frontiera, R. & Mathies, R. A. Direct observation of anharmonic coupling in the time domain with femtosecond stimulated Raman scattering. *Phys. Rev. Lett.* **96**, 238303 (2006).
25. Coe, J. D., Levine, B. G. & Martínez, T. J. Ab initio molecular dynamics of excited-state intramolecular proton transfer using multireference perturbation theory. *J. Phys. Chem. A* **111**, 11302–11310 (2007).
26. Agmon, N. The Grotthuss mechanism. *Chem. Phys. Lett.* **244**, 456–462 (1995).
27. Maddalo, S. L. & Zimmer, M. The role of the protein matrix in green fluorescent protein fluorescence. *Photochem. Photobiol.* **82**, 367–372 (2006).
28. Cerullo, G. *et al.* Photosynthetic light harvesting by carotenoids: detection of an intermediate excited state. *Science* **298**, 2395–2398 (2002).

**Author Contributions** C.F., R.R.F. and R.A.M. designed the research. C.F. and R.R.F. performed the spectroscopic measurements and DFT calculations. C.F. analysed the data and simulated the spectra. R.T. expressed and purified the protein samples. C.F. and R.A.M. wrote the paper. All authors discussed and edited the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.A.M. (ramathies@berkeley.edu).

# METHODS

**Protein preparation.** Plasmids containing the gene for His-tagged wild-type GFP in the pRSETb vector were a gift from R. Wachter (Arizona State University), and were transformed into BL21(DE3)pLysS cells (Novagen). Cultures were grown to an absorbance of $A_{600 nm} \approx 0.5$–0.8 and cooled to 20 °C. A solution of 0.5 mM isopropyl β-D-thiogalactopyranoside (Promega) was then added to induce protein expression, and the cultures were incubated in a shaker for a further 14–16 h. Collected cell pellets were resuspended in buffer A (10 mM Tris, 100 mM NaCl, 50 mM $Na_2HPO_4$, pH 8.0) with 1 mM Pefabloc (Pentapharm), DNase I and lysozyme, and lysed using an Emulsiflex-C5 high-pressure homogenizer at 15,000 psi (Avestin). Lysed cells were centrifuged at 42,000 r.p.m. (200,000g) in an XL-100K ultracentrifuge (Beckman Optima) with a Ti-45 rotor (Beckman Optima) for 1 h and the supernatant was applied to a 30-ml Ni-NTA column (Qiagen). The column was washed with 5–10 column volumes of buffer B (50 mM $KH_2PO_4$, 300 mM NaCl, 20 mM imidazole, 5% glycerol, pH 7.9), and the protein was eluted with 5 column volumes of buffer C (50 mM $KH_2PO_4$, 300 mM NaCl, 250 mM imidazole, 5% glycerol, pH 7.9). To remove the excess imidazole, the purified wild-type GFP was then exchanged into buffer A containing 5% glycerol by spin concentration (Millipore) for storage at −80 °C. The final GFP concentration was adjusted to have an absorbance of $A_{396 nm} = 1$ per millimetre. For the $D_2O$ experiments, GFP was exhaustively exchanged with buffer A in deuterium oxide (99.9% deuterium; Cambridge Isotope Laboratories) and allowed to sit for three days before measurement. The expressed wild-type GFP in our experiments contains the ubiquitous Q80R substitution on the protein surface, which is not known to have any effect on the protein properties[4]. Sample integrity was determined by ultraviolet-visible spectroscopy before and after FSRS.

**FSRS spectroscopy.** FSRS was performed using an apparatus that has previously been described in detail[10]. In brief, the femtosecond laser system consists of a mode-locked Ti:sapphire oscillator and regenerative amplifier (Alpha-1000/US, BMI). The 1-kHz amplified pulse train consists of ~700-μJ, 795-nm, 60-fs pulses with energy stability between −0.3% and 0.3%. Around 10% of the pulse energy was frequency-doubled in a 1-mm-thick BBO crystal to generate the 397-nm actinic pump pulse, which was attenuated to ~400 nJ for electronic excitation of the sample. About 1–2% of the amplifier output was used for continuum generation in a 3-mm-thick sapphire plate and then prism-compressed to produce the ~20-fs, 7-nJ probe pulse at the sample. The remaining amplifier output was spectrally filtered to produce the ~2-μJ, 790–795-nm, 3–4-ps Raman pump pulse. All pulses were horizontally polarized.

To perform FSRS experiments, the three-dimensional position of the 1-mm flow cell was adjusted using the optical-Kerr-effect signal of ethanol between the actinic pump and the Raman probe (non-collinear configuration) to maximize the signal and establish the time zero. The cross-correlation time was 130 fs. The Raman pump pulse was then adjusted to be collinear and temporally overlapped with the Raman probe to generate the maximal stimulated Raman signal of ethanol. The cell was washed three times with $H_2O$ or $D_2O$ before adding ~600 μl of GFP sample solution. The reference pulse and the transmitted probe pulse carrying the stimulated Raman gain features from the sample solution were

spectrally dispersed in a spectrograph (HR320, ISA) with a grating with 600 grooves per millimetre, and recorded using a dual-photodiode array detector (DPDA-1024, Roper Scientific) consisting of two 1,024-pixel arrays of $25 \times 2,500 \mu m^2$ photodiodes vertically separated by 1 mm. The frequency axis was calibrated by collecting the Raman spectrum of cyclohexane, and the spectral resolution was ~15 cm$^{-1}$.

The time delay points between the actinic pump pulse and the subsequent Raman pump-probe time zero were set in random order to improve the accuracy of the actinic pump translation stage, and the ground-state spectrum was measured periodically during the full scan to monitor sample degradation. Each time delay point was an average of at least 60,000 pulses. The buffer spectrum was measured in the same manner immediately after the protein measurement. The complete sets of excited-state spectra in both $H_2O$ and $D_2O$ were reproduced three times ($n = 3$). The ground-state spectrum of the wild-type GFP chromophore was also measured, using a 520-nm, 420-μW, ~450-fs Raman pump in a similar FSRS set-up[29]. Under these conditions, the pre-resonantly enhanced anionic form (B; wavelength of absorption maximum, $\nu_{max} \approx 476 nm$) of the chromophore dominates the Raman spectrum and can be compared to the ground-state spectrum of the neutral form (A).

**Data analysis.** The ground-state spectrum of the wild-type GFP neutral chromophore was obtained by subtracting the buffer-solution Raman spectrum from the protein-solution spectrum using a 794-nm Raman pump. The same procedure was performed with a 520-nm Raman pump to obtain the anionic-form spectrum. The individual time-dependent FSRS excited-state spectra after 396–397-nm excitation were generated after one-to-one subtraction of the ground-state spectrum of A from the raw spectrum measured at each delay time for the actinic pump. In this way, we largely removed systematic noise from the measurement. Owing to the small amount of bleaching of ground-state peaks that are not resonantly enhanced by the ~800-nm Raman pump pulse, there are no significant negative peaks in these difference spectra; therefore, the spectral features originate solely from the excited state of the neutral chromophore (A*). A smooth and featureless baseline was drawn and subtracted from each delay difference spectrum to obtain the pure time-dependent FSRS data for detailed analysis. The Levenberg–Marquardt algorithm was used to perform iterative multi-Gaussian-profile curve fitting of every prominent vibrational band appearing in each excited-state spectrum. The peak width has a standard deviation of less than 10% of its mean value; owing to this small variation, peak area instead of peak amplitude was used to best represent the spectral intensity. The peak centre frequency has a standard deviation of ±1–4 cm$^{-1}$ for the various fitted modes (shown in Fig. 3a as error bars). The low-frequency components of the observed spectral oscillations were retrieved by Fourier-transforming the time-dependent data up to 1 ps. DFT calculations[20] were performed on the ground states of the neutral and anionic forms of the isolated model chromophore *in vacuo* to support vibrational peak assignments.

29. Shim, S. & Mathies, R. A. Generation of narrow-bandwidth picosecond visible pulses from broadband femtosecond pulses for femtosecond stimulated Raman. *Appl. Phys. Lett.* **89,** 121124 (2006).

# LETTERS

# Oxygen and hydrogen isotope evidence for a temperate climate 3.42 billion years ago

M. T. Hren[1]†, M. M. Tice[2] & C. P. Chamberlain[3]

Stable oxygen isotope ratios ($\delta^{18}O$) of Precambrian cherts have been used to establish much of our understanding of the early climate history of Earth[1–3] and suggest that ocean temperatures during the Archaean era (~3.5 billion years ago) were between 55 °C and 85 °C (ref. 2). But, because of uncertainty in the $\delta^{18}O$ of the primitive ocean, there is considerable debate regarding this conclusion. Examination of modern and ancient cherts indicates that another approach, using a combined analysis of $\delta^{18}O$ and hydrogen isotopes ($\delta D$) rather than $\delta^{18}O$ alone, can provide a firmer constraint on formational temperatures without independent knowledge of the isotopic composition of ambient waters[4,5]. Here we show that $\delta^{18}O$ and $\delta D$ sampled from 3.42-billion-year-old Buck Reef Chert rocks in South Africa are consistent with formation from waters at varied low temperatures. The most $^{18}O$-enriched Buck Reef Chert rocks record the lowest diagenetic temperatures and were formed in equilibrium with waters below ~40 °C. Geochemical and sedimentary evidence suggests that the Buck Reef Chert was formed in shallow to deep marine conditions, so our results indicate that the Palaeoarchaean ocean was isotopically depleted relative to the modern ocean and far cooler (≤40 °C) than previously thought[2].

Reconstructions of environmental parameters of the early Earth provide critical boundary conditions for understanding the origin and evolution of life, yet estimates of early climatic temperatures are highly controversial. Geochemical and isotopic compositions of cherts, defined here as rocks composed of >95% microcrystalline quartz, have been used extensively to infer early ocean composition and temperature[1–3]. Cherts from the Onverwacht group of the Palaeoarchaean (3.6–3.2 billion years old) Barberton greenstone belt are significantly depleted in $^{18}O$ relative to recent marine chert[1,2,6] and maximum $\delta^{18}O$ values are similar to the minimum $\delta^{18}O$ of Phanerozoic (~542 million years ago to present) marine cherts[6]. The oxygen isotope composition of chert can be used as a palaeothermometer because the fractionation between $^{18}O$ in silica and ambient water ($\Delta^{18}O = \delta^{18}O_{chert} - \delta^{18}O_{water}$) is dependent on temperature $T$ and can be described by the quartz–water fractionation equation[7,8]: $1,000\ln[\Delta^{18}O] = (3.09 \times 10^6\, T^{-2}) - 3.29$. Low Palaeoarchaean chert $\delta^{18}O$ values have been interpreted as evidence of a hot early ocean (55 °C to 85 °C)[1,2] and cited in support of palaeotemperature reconstructions from protein thermostabilities[9] and $\delta^{30}Si$ isotopes[3].

Inferences of 'hot' Palaeoarchaean climatic temperatures have been controversial[2,3,10], in part because $\delta^{18}O$ determinations of palaeotemperature rely on the assumption that the Archaean ocean $\delta^{18}O$ was similar to that in an ice-free modern ocean[2]. However, oxygen isotope data[11] and models of long-term ocean composition suggest that $\delta^{18}O$ could have varied significantly over time, complicating $\delta^{18}O$ palaeotemperature estimates[10,12,13]. Palaeoweathering proxies have

proved similarly difficult to interpret unambiguously in terms of early climatic temperatures[14,15] and variations in $\delta^{30}Si$ data that were previously interpreted as a record of variable ocean temperature have been reinterpreted as variations in source fluids, not ocean temperature[16].

Paired chert $\delta D$ and $\delta^{18}O$ values can provide constraints on crystallization temperature when the $\delta^{18}O$ of ambient fluid is unknown owing to the shape of the temperature-dependence curve of $^{18}O$ and D fractionation[4,5]. Though chert does not contain hydrogen as a structural component, the quartz phases that comprise it (microcrystalline quartz and chalcedony) contain a minor amount of non-exchangeable, non-stoichiometric hydroxyl groups and hydroxyl-derived water (<1 wt%)[4,17]. The oxygen and hydrogen isotope composition of chert is set during transformations of precursor opaline or silica phases to microcrystalline quartz. At low temperatures, $\delta D_{chert}$ values are up to 80‰ lower than the $\delta D_{water}$ from which cherts formed, while at temperatures greater than ~55 °C, D/H fractionation is similar in magnitude to that of hydrogen in hydrous clay minerals (~20‰ to 30‰)[4]. Cherts formed from recent ocean waters at low temperatures (<55 °C) plot on or below a line with a slope close to −6 (line A in Fig. 1) whereas those formed at higher temperatures (>55 °C) exhibit smaller isotopic fractionations and lie along an array with a slope of less than ~1 (Fig. 1)[4,5]. Isotopic trends consistent with temperature-controlled fractionation are preserved in chert nodules from the Jurassic period[18] and in sediments from the Eocene epoch[19].

We analysed the $\delta D$ and $\delta^{18}O$ of 27 black-and-white banded cherts from the 3,416 ± 5-million-year-old Buck Reef Chert of the Onverwacht formation in South Africa (see Supplementary Information) to provide new constraints on Palaeoarchaean ocean temperatures. Distributions of sedimentary structures, deformed sedimentary layers, grain types and sizes, and lamination styles indicate that rocks in the studied section were formed from sediments deposited under increasing water depth, from shallow coastal lagoons to a basinal setting[20,21]. In shallow-water environments, rapid precipitation of isopachous silica cements, probably within one metre of the sediment–water interface, prevented compaction of abundant carbonaceous grains[21–23]. Open voids in these bands subsequently filled with higher-order silica and quartz precipitates. Black chert bands in deep-water environments were silicified after compaction. In all environments, sediment disruption by storms produced deformed white bands as plastic or brittle solids within a fluid black matrix, indicating that white chert bands silicified before black chert bands[21–23]. White bands preserve no relict textures of primary precipitates other than opal, suggesting that they did not undergo prolonged silicification. All silica phases have subsequently been converted to microcrystalline quartz, and subjected to peak metamorphic temperatures of ~315 °C (refs 21, 24).

Empirical and theoretical chert–water fractionation data show that cherts formed at low temperatures (<55 °C) lie in a $\delta D$–$\delta^{18}O$ array

[1]Department of Geology and Geophysics, Yale University, 210 Whitney Avenue, New Haven, Connecticut 06511, USA. [2]Department of Geology and Geophysics, Texas A&M University, Room 316 Halbouty, MS 3115, College Station, Texas 77843, USA. [3]Department of Environmental Earth System Science, Stanford University, 450 Serra Mall, Braun Hall, Building 320, Stanford, California 94305, USA. †Present address: Department of Geological Sciences, University of Michigan, 2534 C.C. Little Building, 1100 North University Avenue, Ann Arbor, Michigan 48109-1005 (M.T.H.).
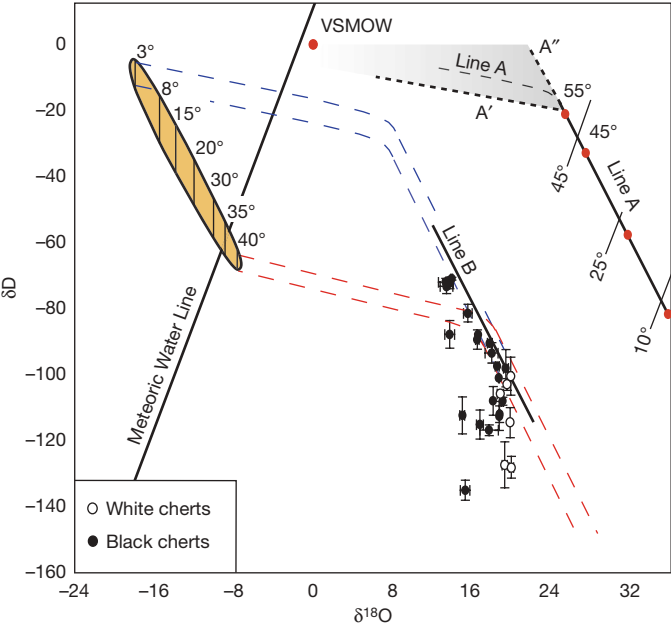
**Figure 1 | δ$^{18}$O and δD of Buck Reef Cherts.** Line A and the grey shaded region A′–A′′ provide the range of temperature-dependent fractionation paths for chert in equilibrium with VSMOW. Buck Reef Chert rocks are bounded by a line (line B) parallel to recent marine cherts formed at low temperatures (line A; red circles show the expected composition of recent marine cherts formed at selected temperatures from 10–55 °C). Data significantly below Line B reflects post-depositional alteration by isotopically depleted fluids at <250 °C. The isotopic composition of the Palaeoarchaean ocean (yellow shaded ellipse) is constrained by ±5‰ in the intercept of temperature-dependent line B (blue and red dashed lines) and high-temperature (>55 °C) fractionation bounds A′–A′′ (see Supplementary Information). The highest δ$^{18}$O$_{chert}$ value (+20.0‰) records the lowest formation temperature and provides an upper limit on Palaeoarchaean ocean temperature. Formation temperatures are determined for the range of possible ocean compositions and shown as vertical isotherms. Cherts formed near the seawater–sediment interface record temperatures below ~40 °C and a D- and $^{18}$O-depleted Palaeoarchaean ocean. Error bars reflect 1σ uncertainties. The Meteoric Water Line is shown for reference.
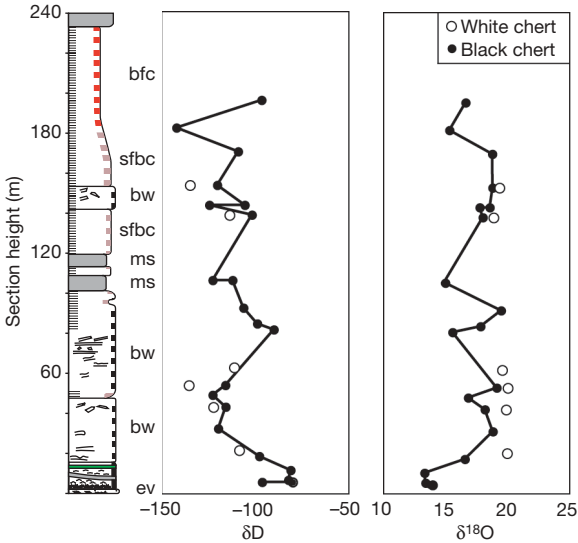


**Figure 2 | Buck Reef Chert stratigraphy.** δ$^{18}$O and δD of black and white cherts from the Buck Reef Chert unit. Large isotopic variations are preserved throughout the sedimentary column. 'ev' represents black cherts deposited in shallow coastal lagoons, 'bw' black-and-white banded cherts deposited on a storm-active platform, 'ms' mafic sills, 'sfbc' slightly ferruginous banded chert deposited on a low-energy platform/basin, 'bfc' banded ferruginous chert deposited in a low-energy basin.

with a steep negative slope, while those formed at higher temperatures (>55 °C) are characterized by relatively large variations in δ$^{18}$O but small differences in δD (refs 4, 5). Post-depositional alteration or contributions of non-hydroxyl hydrogen during analysis of chert δD by high-temperature pyrolysis can produce δ$^{18}$O or δD values that fall below the expected temperature-dependent fractionation line. Measured δ$^{18}$O and δD values for Buck Reef Chert rocks range from +13.7‰ to +20.0‰ and −140‰ to −72‰, respectively (Table 1). δ$^{18}$O values are similar to previous isotope measurements of Barberton cherts[1,2] and both δ$^{18}$O and δD vary widely across the stratigraphic section (Fig. 2). δ$^{18}$O and δD values for black and white chert samples co-vary, with most falling along an array parallel to the

**Table 1 | Isotopic values, carbon and water contents for chert samples**

| Sample | δD | s.d. | n | δ$^{18}$O | s.d. | n | Water (wt%) | δ$^{13}$C | Carbon (wt%) before treatment | Carbon (wt%) after treatment |
|---|---|---|---|---|---|---|---|---|---|---|
| TSA5-1A | −72 | 1.8 | 4 | 13.7 | 0.3 | 3 | 0.38 | −28.6 | 0.20 | 0.00 |
| TSA5-1b | −88 | 4.2 | 3 | 13.7 | 0.5 | 2 | 0.46 | n.m. | 0.00 | 0.00 |
| TSA5-1c | −70 | 0.4 | 3 | 14.0 | 0.1 | 2 | 0.14 | n.m. | 0.00 | 0.00 |
| TSA5-2 | −73 | 2.2 | 3 | 13.4 | 0.6 | 2 | 0.17 | −31.7 | 0.41 | 0.10 |
| TSA5-3 | −72 | 1.2 | 3 | 13.3 | 0.7 | 2 | 0.31 | −35.1 | 0.68 | 0.23 |
| TSA5-4* | −100 | 5.8 | 2 | 19.9 | 0.1 | 2 | 0.08 | −26.2 | 0.13 | 0.02 |
| TSA5-7A* | −115 | 4.5 | 2 | 19.8 | n.a. | 1 | 0.10 | n.m.s. | 0.45 | 0.05 |
| TSA5-7B | −108 | 4.3 | 3 | 18.1 | n.a. | 1 | 0.08 | −31.3 | 0.24 | 0.05 |
| TSA5-8* | −103 | 1.7 | 2 | 19.5 | 0.3 | 4 | 0.13 | −19.9 | 0.10 | 0.02 |
| TSA5-9A* | −128 | 3.3 | 3 | 20.0 | 0.2 | 2 | 0.19 | −24.4 | 0.25 | 0.12 |
| TSA5-9B | −108 | 1.2 | 3 | 19.1 | 0.1 | 2 | 0.08 | n.m. | 0.05 | 0.00 |
| TSA5-10 | −112 | 2.5 | 3 | 18.8 | 0.0 | 3 | 0.10 | n.m. | 0.04 | 0.00 |
| TSA5-11 | −81 | 2.7 | 2 | 15.6 | 0.4 | 2 | 0.25 | −30.6 | 0.38 | 0.09 |
| TSA5-13 | −98 | 5.7 | 2 | 19.4 | 0.2 | 2 | 0.08 | −22.7 | 0.15 | 0.03 |
| TSA5-17A* | −106 | 3.9 | 2 | 18.8 | 0.3 | 2 | 0.08 | −25.1 | 0.16 | 0.07 |
| TSA5-17B | −93 | 3.1 | 3 | 18.0 | 0.7 | 2 | 0.18 | n.m. | 0.00 | 0.00 |
| TSA5-18A | −97 | 2.3 | 2 | 18.5 | 0.2 | 3 | 0.07 | n.m. | 0.05 | 0.03 |
| TSA5-18B | −117 | 1.6 | 2 | 17.7 | 0.9 | 3 | 0.11 | −29.4 | 0.26 | 0.06 |
| TSA5-20 | −88 | 3.4 | 2 | 16.6 | n.a. | 1 | 0.13 | −24.9 | 0.32 | 0.10 |
| TSA5-23 | −140 | 3.0 | 3 | 15.3 | 0.5 | 3 | 0.28 | n.m.s. | n.m.s. | 0.17 |
| TSA5-29 | −115 | 4.4 | 3 | 16.8 | 0.4 | 2 | 0.43 | −22.8 | 0.27 | 0.21 |
| TSA5-32 | −112 | 5.6 | 4 | 15.0 | 0.3 | 3 | 0.10 | −25.1 | 0.23 | 0.00 |
| TSA5-34A* | −127 | 7.0 | 4 | 19.3 | 0.3 | 2 | 0.11 | −25.9 | 0.26 | 0.07 |
| TSA5-34B | −112 | 4.4 | 2 | 18.7 | n.a. | 1 | 0.25 | n.m. | 0.00 | 0.00 |
| TSA5-35B | −101 | 3.0 | 2 | 18.7 | 0.0 | 2 | 0.10 | −22.9 | 0.10 | 0.03 |
| TSA5-40 | −90 | 2.4 | 3 | 17.8 | 0.2 | 2 | 0.07 | −31.7 | 0.32 | 0.04 |
| SAF475-12 | −89 | 2.9 | 3 | 16.5 | 0.2 | 3 | 0.10 | −30.8 | 0.25 | 0.10 |

s.d., Standard deviation. n, Number of samples. n.m., Not measurable, because the carbon content was too low to produce reliable δ$^{13}$C measurements. n.m.s., Not measurable, because of sample issues or available materials. n.a., Not applicable.
* White chert.

line defining the isotopic composition of geologically recent marine cherts formed at low temperatures (<55 °C; Fig. 1)[1,5]. To determine whether the isotopic array records temperature-dependent fractionation processes related to initial diagenetic transformation and subsequent alteration of some primary chert $\delta^{18}O$ or $\delta D$ values, we used Monte Carlo simulations to define the slopes of the bounding lines to the Buck Reef Chert $\delta^{18}O$ and $\delta D$ data (line B) and the original [18]O and D fractionation calibration data set[4] (Supplementary Information). Simulation results show that the calculated slope of the upper bound to the Buck Reef Chert data ($-4.8^{+0.6}_{-1.1}$; 95% confidence) is indistinguishable from the proposed slope of the low-temperature [18]O and D fractionation line ($-5.9^{+1.0}_{-0.6}$; 95% confidence) at the 95% confidence level ($P = 0.216$). Thus, it is probable that the cherts that form the array bounded by line B reflect temperature-dependent fractionation from a single fluid composition and rocks with the most enriched $\delta^{18}O$ could only have been produced by crystallization at moderate to cold temperatures (<55 °C). If cherts that are bounded by the temperature-dependent line B (Fig. 1) preserve primary $\delta^{18}O$ and $\delta D$, this array can constrain the temperature and isotopic composition of the Palaeoarchaean ocean.

The determination of whether cherts bounded by line B retain primary $\delta^{18}O$ and $\delta D$ is critical to relating calculated formation temperatures to Palaeoarchaean ocean conditions. Recent studies suggest that volcanics underlying the Buck Reef Chert unit were silicified after emplacement by low-temperature (<150 °C) hydrothermal alteration[25]. Post-depositional alteration of Buck Reef Chert rocks could reset some or all chert $\delta^{18}O$ or $\delta D$ values. Several lines of evidence have been used to argue against pervasive resetting or overprinting of chert $\delta^{18}O$, including preservation of large variations in $\delta^{18}O$ across stratigraphic, regional and metamorphic gradients[1,2], similar magnitudes of [18]O depletion (6–10‰) in Archaean chert and carbonate[1,2,11], and correlations between $\delta^{30}Si$ and $\delta^{18}O$ (ref. 3). However, despite evidence for preservation of primary chert $\delta^{18}O$, there is considerable uncertainty over the preservation potential of original $\delta D$.

There are two important concerns regarding preservation of chert $\delta D$. First, the water contents of Buck Reef Chert rocks are significantly lower than modern opaline silica[4,5], suggesting possible hydrogen loss due to dehydration during metamorphism. Experimental fractionation studies show that dehydration of hydrous minerals can result in progressive D-depletion of mineral-bound hydrogen[26]. However, metamorphic dehydration of hydrous minerals produces a clear relationship between $\delta D$ and weight per cent water in the mineral reservoir but chert data show no such relationship, which suggests that mineral dehydration is not the primary cause of low $\delta D$ values (Supplementary Information). A second concern is that pervasive post-depositional recrystallization could have altered the original isotopic compositions. The potential impact of hydrothermal alteration on primary chert $\delta^{18}O$ and $\delta D$ can be examined by calculating isotope exchange paths during water–rock interaction for varied temperatures and fluid compositions.

The degree of isotopic fractionation associated with the equilibration of waters with host rock is dependent upon the temperature and water–rock mole ratio[27]. For a system in equilibrium between water and rock, the effect of variable water–rock ratios on the isotopic composition of exchanged water or rock can be described by the equation: $W/R = (\delta_{rock}^{final} - \delta_{rock}^{initial})/(\delta_{water}^{initial} - (\delta_{rock}^{final} - \Delta_{rock-water}))$, where $W$ is the atom per cent water oxygen in the total system, $R$ is the atom per cent exchangeable rock oxygen in the bulk system, and $\Delta$ is the rock–water fractionation for either hydrogen or oxygen[27].

Figure 3 shows the calculated isothermal exchange path for a hypothetical isotopically depleted water ($\delta^{18}O = -14$‰, $\delta D = -102$‰) equilibrating with chert of similar composition to Buck Reef Chert samples at the low end of the isotope array ($\delta^{18}O = +20$‰, $\delta D = -100$‰), which are cherts interpreted to have formed directly from sea water[22]. The effect of equilibration with isotopically depleted waters at hydrothermal or low metamorphic temperatures (150 °C to
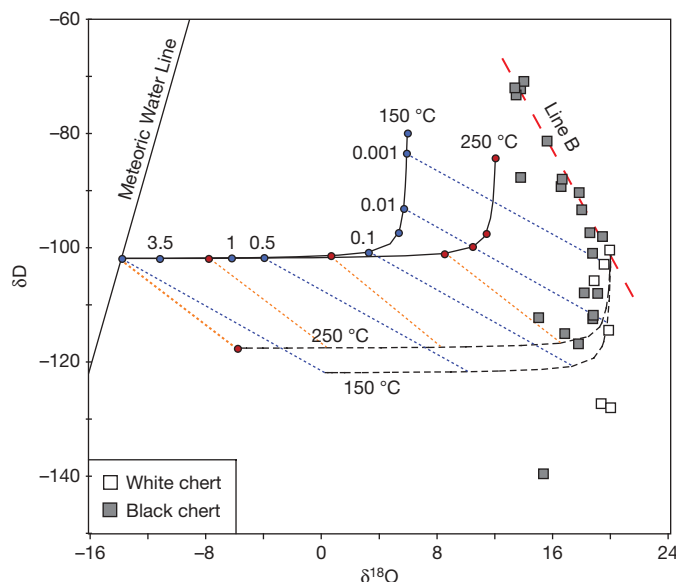


**Figure 3 | Isothermal exchange paths during water–rock interaction.** Calculated isotopic evolution of isotopically depleted hydrothermal fluids (−14‰ and −102‰ for $\delta^{18}O$ and $\delta D$, respectively) and chert with a composition similar to the most ^{18}O-enriched white chert ($\delta^{18}O = +20.0$‰; $\delta D = -100$‰) during water–rock interaction at 150 °C and 250 °C. Exchange paths for fluids (solid lines) and chert (dashed lines) are calculated using the temperature-dependent quartz–water [18]O fractionation factor[8] and a theoretical D/H fractionation factor that assumes high-temperature chert–water D/H fractionation is similar to clays[4,5] and is based on kaolinite–water D/H fractionation. Water and corresponding chert isotopic compositions for varied water–rock ratios are shown by blue and orange dashed lines.

250 °C) and low water–rock ratios is an increased vertical spread ($\delta D$) of the data, similar to what we observed. At high water–rock ratios, equilibration with isotopically depleted waters is expected to increase the spread of $\delta^{18}O$ values. The white chert bands sampled have six of the nine highest $\delta^{18}O$ values measured, and a much lower sample variance than the black bands (0.2 versus 4.4, $P = 0.001$ for the $F$ test). This result is consistent with the hypothesis that white bands precipitated both before black bands and over a smaller range of burial depths, and also suggests that primary $\delta^{18}O$ values have not been substantially reset or overprinted. In contrast, white bands exhibit no statistically significant difference in $\delta D$ sample variance from black bands (150 versus 330, $P = 0.19$ for the $F$ test). These data suggest that cherts that fall below the bounding line B originally recorded temperature-dependent fractionation, but have been altered by isotopically depleted waters at a range of temperatures and low water–rock ratios (0.001 to 0.1) after deposition. Samples along the temperature-dependent array most probably represent primary chert isotope compositions and record Palaeoarchaean ocean temperature and isotopic composition.

Calculation of chert formation temperature from $\delta^{18}O$ requires an estimate of $\delta^{18}O_{water}$. The isotopic composition of ambient fluids at the time of formation can be constrained using the low-temperature (<55 °C) isotope fractionation (line A) with a ±5‰ uncertainty in the intercept of the bounding line B ($\delta D = -6 \times \delta^{18}O + 16.5$), and the assumption that high-temperature D-fractionation is similar to hydrogen in clay minerals[4,5] (line A′ in Fig. 1), or a less likely scenario that D-fractionation is reduced to zero following the low-temperature fractionation line A to 75 °C, above which there is no D-fractionation between water and chert (line A″ in Fig. 1). Assuming all Buck Reef Chert rocks bounded by line B fall on or below a fractionation line similar to line A and within the bounds defined by A′–A″, the isotopic composition of chert-forming waters can be constrained to $-18$‰ < $\delta^{18}O$ < $-8$‰ and $-70$‰ < $\delta D$ < $-5$‰ (Fig. 1; see Supplementary Information for details).

**207**

The temperature of chert formation is calculated using the quartz–water oxygen isotope fractionation factor and an estimate of water $\delta^{18}O$. White cherts that are inferred to have formed closest to the seawater–sediment interface have the highest $\delta^{18}O$ and lowest $\delta D$ along the temperature-dependent line B and produce the lowest temperatures of the sample suite. These are interpreted to reflect very early diagenetic temperatures, potentially close to those of sea water, while higher temperatures associated with black bands are reflective of prolonged diagenetic silicification during deeper burial. Over the possible range of fluid $\delta^{18}O$ compositions, the maximum temperature calculated for white chert bands ($\delta^{18}O = +20.0‰$) is ~40 °C, suggesting that the Buck Reef Chert ocean could have been no hotter. These data are thus incompatible with a 'hot' Archaean ocean, and are consistent with a marine origin for these rocks, as opposed to direct hydrothermal precipitation[21,28].

Models of the $\delta^{18}O$ evolution of sea water predict that Archaean ocean water may have been depleted in $^{18}O$ by as much as $-13.3‰$ (refs 10, 29). If the $\delta^{18}O$ of the Archaean ocean were $-10‰$, as suggested by one model[10], the calculated temperatures of formation for all Buck Reef Chert samples fall between 32 °C and 65 °C. Significantly, this interpretation would also imply a radically deuterium-depleted early ocean ($\delta D = -60‰$), an outcome consistent with recent models that suggest significant D-depletion of the early ocean reservoir and a prolonged period of hydrogen escape from the atmosphere[30].

## METHODS SUMMARY

Chert pieces were powdered and sieved through 170 mesh and treated in weak HCl for 10 min to remove carbonate. Rinsed powders were placed in a weak sodium hypochlorite solution for 24 h to remove organic matter, and rinsed to remove residual NaOCl. Comparison of treated and untreated samples shows near-complete removal of organics in most samples with no impact on measured isotopic compositions in 'pristine' hand samples. $\delta^{18}O$ was measured by infrared laser fluorination of 0.8 to 1.2 mg of chert powder using $BrF_5$ as an oxidizing agent. $O_2$ gas was analysed on a Finnigan MAT 252 mass spectrometer with a precision of $\pm0.2‰$. Chert $\delta D$ measurements were made by high-temperature pyrolysis. 10–20 mg of treated chert powder was dried in a vacuum oven at 80 °C for 2 h to remove adsorbed water, then transferred to a He-flushed autosampler. $H_2$ gas was measured on a Finnigan Delta Plus XL in continuous flow. Standard analyses yield a precision of $\pm2.5‰$. $\delta^{18}O$ and $\delta D$ data are reported relative to Vienna Standard Mean Ocean Water (VSMOW). Carbon isotope measurements were made by high-temperature combustion in continuous flow with He as carrier gas and attached to a Finnigan Delta Plus XL. Data are reported relative to Vienna Pee Dee Belemnite (VPDB). Precision of $\delta^{13}C$ analyses is $\pm0.2‰$. The range of isotope compositions of chert formation waters is constrained by the condition that cherts along line B fall on or below a line with a slope equal to the temperature-dependent fractionation line A and within bounds defined by line A'–A''. Temperatures of formation for white chert inferred to have formed close to the seawater–sediment interface ($\delta^{18}O = +20.0‰$) were calculated using the quartz–water oxygen isotope fractionation factor[7] and the range of possible fluid $\delta^{18}O$ compositions determined in Fig. 1.

1. Knauth, L. P. & Lowe, D. R. Oxygen isotope geochemistry of cherts from the Onverwacht group (3.4 Ga), Transvaal, South Africa, with implications for secular variations in the isotopic compositions of cherts. *Earth Planet. Sci. Lett.* **41**, 209–222 (1978).
2. Knauth, L. P. & Lowe, D. R. High Archean climatic temperature inferred from oxygen isotope geochemistry of cherts in the 3.5 Ga Swaziland Supergroup, South Africa. *Geol. Soc. Am. Bull.* **115**, 566–580 (2003).
3. Robert, F. & Chaussidon, M. A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. *Nature* **443**, 969–972 (2006).
4. Knauth, L. P. & Epstein, S. Hydrogen and oxygen isotope ratios in nodular and bedded cherts. *Geochim. Cosmochim. Acta* **40**, 1095–1108 (1976).
5. Knauth, L. P. *Oxygen and Hydrogen Isotope Ratios in Cherts and Related Rocks.* 1–369, PhD thesis (CalTech, 1973).
6. Perry, E. C. Jr. The oxygen isotope chemistry of ancient cherts. *Earth Planet. Sci. Lett.* **3**, 62–66 (1967).
7. Clayton, R. N., O'Neil, J. R. & Mayeda, T. K. Oxygen isotope exchange between quartz and water. *J. Geophys. Res.* **77**, 3057–3067 (1972).
8. Knauth, L. P. & Epstein, S. Hydrogen and oxygen isotope ratios in silica from the JOIDES Deep Sea Drilling Project. *Earth Planet. Sci. Lett.* **25**, 1–10 (1975).
9. Gaucher, E. A., Govindarajan, S. & Ganesh, O. K. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* **451**, 704–708 (2008).
10. Kasting, J. F. *et al.* Paleoclimates, ocean depth, and the oxygen isotopic composition of seawater. *Earth Planet. Sci. Lett.* **252**, 82–93 (2006).
11. Shields, G. & Vezier, J. Precambrian marine carbon isotope database: version 1. *Geochem. Geophys. Geosyst.* **3**, doi:10,1029/2001GC000266 (2002).
12. Walker, J. C. G. & Lohmann, K. C. Why the oxygen isotopic composition of seawater changes with time. *Geophys. Res. Lett.* **16**, 323–326 (1989).
13. Wallman, K. Impact of atmospheric $CO_2$ and galactic cosmic radiation on Phanerozoic climate change and the marine $\delta^{18}O$ record. *Geochem. Geophys. Geosyst.* **5**, doi:10.10029/2003G000683 (2004).
14. Sleep, N. H. & Hessler, A. M. Weathering of quartz as an Archean climatic indicator. *Earth Planet. Sci. Lett.* **241**, 594–602 (2006).
15. Lowe, D. R. A comment on "Weathering of quartz as an Archean climatic indicator" by N.H. Sleep and A.M. Hessler. *Earth Planet. Sci. Lett.* **253**, 530–533 (2007).
16. van den Boorn, S. H. J. M., van Bergen, M. J., Nijman, W. & Vroon, P. Z. Dual role of seawater and hydrothermal fluids in Early Archean chert formation: evidence from silicon isotopes. *Geology* **35**, 939–942 (2007).
17. Knauth, L. P. & Epstein, S. The nature of water in hydrous silica. *Am. Mineral.* **67**, 510–520 (1982).
18. Sharp, Z. D., Durakiewicz, T., Migaszewski, Z. M. & Atudorei, V. N. Antiphase hydrogen and oxygen isotope periodicity in chert nodules. *Geochim. Cosmochim. Acta* **66**, 2865–2873 (2002).
19. Abruzzese, M. J., Waldbauer, J. R. & Chamberlain, C. P. Oxygen and hydrogen isotope ratios in freshwater chert as indicators of ancient climate and hydrologic regime. *Geochim. Cosmochim. Acta* **69**, 1377–1390 (2005).
20. Tice, M. M. & Lowe, D. R. Photosynthetic microbial mats in the 3,416-Myr-old ocean. *Nature* **431**, 549–552 (2004).
21. Tice, M. M. & Lowe, D. R. The origin of carbonaceous matter in pre-3.0 Ga greenstone terrains: a review and new evidence from the 3.42 Ga Buck Reef Chert. *Earth Planet. Sci. Lett.* **76**, 259–300 (2006).
22. Lowe, D. R. & Knauth, L. P. Sedimentology of the Onverwacht Group (3.4 billion years), Transvaal, South Africa, and its bearing on the characteristics and evolution of the Early Earth. *J. Geol.* **85**, 699–723 (1977).
23. Lowe, D. R. in *Geologic Evolution of the Barberton Greenstone Belt, South Africa* (eds Lowe, D. R. & Byerly, G. R.) Vol. 329, 83–114 (Geological Society of America Special Paper, 1999).
24. Xie, X., Byerly, G. R. & Ferrell, R. E. Jr. IIb trioctahedral chlorite from the Barberton greenstone belt: crystal structure and rock composition constraints with implications to geothermometry. *Contrib. Mineral. Petrol.* **126**, 275–291 (1997).
25. Hofmann, A. & Harris, C. Silica alteration zones in the Barberton greenstone belt: A window into subseafloor processes 3.5–3.3 Ga ago. *Chem. Geol.* **257**, 221–239 (2008).
26. Suzuoki, T. & Epstein, S. Hydrogen isotope fractionation between OH-bearing silicate minerals and water. *Geochim. Cosmochim. Acta* **40**, 1229–1240 (1976).
27. Taylor, H. P. The application of oxygen and hydrogen isotope studies to problems of hydrothermal alteration and ore deposition. *Econ. Geol.* **69**, 843–883 (1974).
28. Tice, M. M. & Lowe, D. R. Hydrogen-based carbon fixation in the earliest-known photosynthetic organisms. *Geology* **34**, 37–40 (2006).
29. Jaffrés, J. B. D., Shields, G. A. & Wallmann, K. The oxygen isotope evolution of seawater: a critical review of a long-standing controversy and an improved geologic water cycle model for the past 3.4 billion years. *Earth Sci. Rev.* **83**, 83–122 (2007).
30. Genda, H. & Ikoma, M. Origin of the ocean on Earth: early evolution of water D/H in a hydrogen-rich atmosphere. *Icarus* **194**, 42–52 (2008).

# Permeability of asthenospheric mantle and melt extraction rates at mid-ocean ridges

James A. D. Connolly[1], Max W. Schmidt[1], Giulio Solferino[1] & Nikolai Bagdassarov[2]

**Magmatic production on Earth is dominated by asthenospheric melts of basaltic composition that have mostly erupted at mid-ocean ridges. The timescale for segregation and transport of these melts, which are ultimately responsible for formation of the Earth's crust, is critically dependent on the permeability of the partly molten asthenospheric mantle, yet this permeability is known mainly from semi-empirical and analogue models[1–6]. Here we use a high-pressure, high-temperature centrifuge, at accelerations of 400*g*–700*g*, to measure the rate of basalt melt flow in olivine aggregates with porosities of 5–12 per cent. The resulting permeabilities are consistent with a microscopic model in which melt is completely connected, and are one to two orders of magnitude larger than predicted by current parameterizations[4,7]. Extrapolation of the measurements to conditions characteristic[8] of asthenosphere below mid-ocean ridges yields proportionally higher transport speeds. Application of these results in a model[9] of porous-media channelling instabilities[10] yields melt transport times of ~1–2.5 kyr across the entire asthenosphere, which is sufficient to preserve the observed $^{230}$Th excess of mid-ocean-ridge basalts and the mantle signatures of even shorter-lived isotopes such as $^{226}$Ra (refs 5,11–14).**

The presence of melt in the mantle below mid-ocean ridges (MORs) has been demonstrated using magnetotelluric and seismic reflection sensing[8], the latter suggesting a minimum of 1–2 vol.% melt. Simple flow models that account for mantle flow, the extent of the melting region and the rate of igneous oceanic crust production, place an upper bound on the average sub-MOR melt fraction of 2–4 vol.% (ref. 15). This estimate has been corroborated by melt inclusion studies, which indicate melt fractions of a few per cent in the solid matrix beneath ocean ridges[16,17]. Melt fraction and extraction rates directly control the composition of MOR basalts in terms of strongly incompatible elements and U-series nuclides[7,18]. Particularly at MORs, the preservation of short-lived radiogenic isotopic signatures (of protactinium, radium, thorium and uranium[7]) requires melt flow rates of 10–100 m yr$^{-1}$ or greater if these signatures reflect the onset of melting at depths in excess of 60 km (ref. 8). Such flow rates are too high to be explained by present models for compaction-driven melt expulsion[7,12,13], a discrepancy that has motivated increasingly complex models of melt transport[9,11,19,20]. In this context, permeability is the most uncertain physical parameter controlling the velocity of the melt relative to the solid residuum, that is, the melt segregation velocity and the timescale of melt generation in the mantle.

In this study, we make direct experimental observations of the rate of gravitationally induced flow of basaltic melts through an olivine matrix as a function of matrix porosity. For melts and melt fractions of interest, rates of self-induced compaction-driven flow are insignificant under normal gravitational acceleration in experimental systems. Accordingly, here we use a piston–cylinder device, which permits us to emulate asthenospheric melting conditions, mounted in a centrifuge[21] to enhance melt flow rates.

The experimental charges were mixtures of MOR basalt glass and San Carlos olivine with an average grain size of ~1.8 μm (Methods and Supplementary Table 1) loaded in cylindrical capsules of 2-mm inner diameter. Before centrifuging, the samples were annealed and texturally equilibrated for 24 h in a standard piston–cylinder apparatus at 1,270–1,300 °C and 1.0 GPa (Methods). The annealed samples were loaded into the centrifuging piston–cylinder, whose axis of radial symmetry (the sample axis) is parallel to the centripetal force of the centrifuge. Discounting minor effects due to drag along the capsule walls and the slight shear deformation of the sample on loading, this geometry ensures one-dimensional compaction-driven melt flow along the sample axis as a consequence of centrifuging. Operating conditions for the centrifuge experiments were as follows: *g* force, 400*g*–700*g*; temperature, 1,270–1,300 °C; pressure, 0.9–1.4 GPa; duration, 4–11 h. Melt distributions in the samples were measured from electron backscatter microscopy images of polished sections cut along the sample axis. Each image was divided into five to seven subsections, orthogonal to the sample axis, within which the proportion of melt was determined by image analysis (Fig. 1).

Comparison of the melt distribution in an uncentrifuged sample with the distributions in the centrifuged samples (Fig. 2) demonstrates that centrifuging significantly enhances melt flow and compaction. Neglecting the low solid and melt compressibility, for one-dimensional flow starting from a uniform porosity distribution, the volumetric barycentre of a sample (Fig. 1) is the point at which the porosity is equal to the initial porosity, $\phi_0$. Assuming Darcyian flow and a constant pressure gradient at the barycentre, and discounting the effect of grain growth, it follows that the melt velocity relative to the barycentre, $\mathbf{v}_0$, is also constant. Thus

$$\mathbf{v}_0 = \frac{\mathbf{u}_z}{\phi_0 t} \int_0^H (\phi - \phi_0)\, \mathrm{d}z \qquad (1)$$

where $\phi$ is the melt-filled porosity, $t$ is the duration of the experiment, $\mathbf{u}_z$ is an upward-directed unit vector and $H$ is the $z$ coordinate at the top of the capsule relative to the sample barycentre (that is, half its height). To evaluate the integral in equation (1), the measured porosities were fitted as a linear function of $z$ (Fig. 2), yielding the melt velocities given in Table 1. Upward melt flow is compensated by downward solid flow, that is, compaction of the olivine matrix. Although this process is usually considered to be viscous[2], the observed compaction rates are at least four orders of magnitude greater than those measured experimentally[22] (Methods) for similar materials and physical conditions. These latter rates are consistent with experimental constraints on olivine shear viscosity; thus, we conclude that compaction in our samples was accomplished by a plastic mechanism induced by centrifuging.

[1]Institute for Mineralogy and Petrology, ETH Zurich, 8092 Zurich, Switzerland. [2]Institut für Geowissenschaften, Facheinheit Geophysik, J. W. Goethe Universität, Altenhöferallee 1, 60438 Frankfurt am Main, Germany.

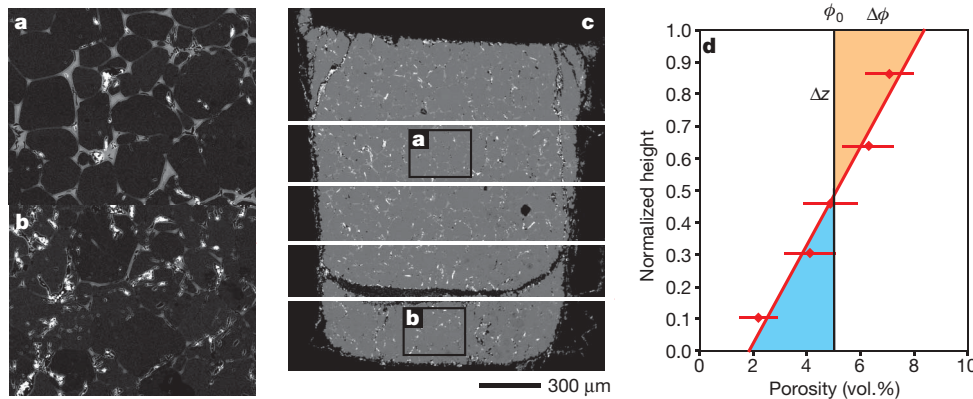**Figure 1 | Centrifuged high-pressure sample and principle of obtaining melt velocities. a–c,** Electron backscatter image of an experimental run product: **a** and **b** are details of **c** demonstrating the increase of melt fraction towards the gravitationally defined 'top' of the sample. The grey interstitial phase is melt and the dark granular phase is olivine. The centrifugal acceleration acts downwards. **d,** The average melt content for each slice is fitted as a linear function of height; the intersection with the average porosity defines the barycentre of the sample, at which the porosity corresponds to the average porosity, $\phi_0$, and remains constant during the experiment. Error bars, $2\sigma$. The blue and orange areas represent the melt volume that has moved through the barycentre plane, and are used to calculate melt velocities according to equation (1).

Compaction-driven flow of a low-viscosity melt through a creeping matrix has two limiting regimes: a hydraulically limited regime (in which the rate of fluid flow is controlled by the matrix permeability and other hydraulic properties) and a rheologically limited regime (in which the rate of melt flow is controlled by the matrix rheology). To extract the permeability of our samples, we assume melt flow to be hydraulically limited, in which case the hydraulic gradient for melt flow is $\sim -\Delta\rho\mathbf{a}$, where $\Delta\rho$ is the melt–olivine density difference ($\sim 400\,\mathrm{kg\,m^{-3}}$) and $\mathbf{a}$ is the acceleration. According to Darcy's law (Methods), the permeability of the olivine matrix is then

$$k = \frac{\mu|\mathbf{v}_0|\phi_0}{\Delta\rho|\mathbf{a}|} \qquad (2)$$

where $\mu$ is the melt viscosity (71 Pa s in the experiments[23]). Applying equation (2) to the experimental results yields permeabilities in the range of $(2–3)\times10^{-14}\,\mathrm{m^2}$ for porosities of 5–12% (Table 1). The symmetry of the experimental porosity distributions suggests rheologically rather than hydraulically limited compaction, in which case equation (2) provides a lower bound for the sample permeabilities.
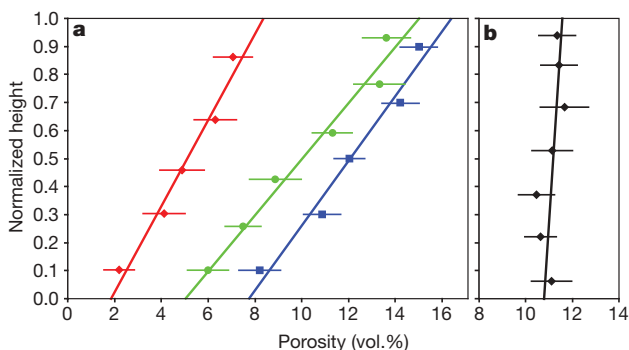


**Figure 2 | Profiles of porosity versus sample height for the basaltic melt centrifuged from the olivine matrix. a,** The three centrifuged experiments: ZOFB-1 (red), experiment run for 4 h at $700g$ with $\phi_0 = 0.051$; ZOFB-3 (green), 4 h at $600g$ with $\phi_0 = 0.097$; ZOFB-2 (blue), 10.5 h at $400g$ with $\phi_0 = 0.121$. **b,** The static experiment (OFB-5), in which the porosity gradient is not statistically significant. Error bars, $2\sigma$. Although the best fit to the static melt distribution has a slight slope $\mathrm{d}\phi/\mathrm{d}z$, a vertically homogeneous melt distribution can be fitted within the errors. The integral of equation (1) is thus formulated with $\phi_0$ as the reference melt distribution along the sample axis. We note that the static experiment also had an annealing time three times longer than in ZOFB-3, the longest-centrifuged experiment, allowing for much longer thermal melt migration.

Models of asthenospheric melt flow generally presume that permeability can be expressed as a function of porosity and grain size within the solid aggregate[2]:

$$k = \frac{d^2\phi^n}{C} \qquad (3)$$

where $d$ is grain size and $C$ and $n$ are parameters that depend on the geometry of the microscopic hydraulic bonds. A quadratic porosity dependence in equation (3) (that is, $n = 2$) is appropriate when melt first exceeds the percolation threshold and becomes interconnected along grain edges[3,5]. At the higher porosities of our experiments (Fig. 1), this dependence becomes cubic[4–7] ($n = 3$) and estimates for the constant $C$ range from 50 to 1,000 (ref. 2), with more recent estimates converging towards values of 200–300 (refs 4,5,7). In contrast, our observations yield $C$ values in the range of 3–27, suggesting that for a given porosity, permeabilities are up to two orders of magnitude greater than previously supposed. Within measurement error, our values for $C$ are constant, but the error does not account for the possibility that the hydraulic gradient during rheologically limited compaction may be much less than $\Delta\rho\mathbf{a}$.

Reference [22] reports an experimental study of forced compaction-driven flow of lithium silicate, basaltic and albitic melts through an olivine matrix. Although we suggest that the experimental results of ref. 22 do not constrain the matrix permeability with respect to the basaltic and lithium silicate melts (Supplementary Information), the experiment on albite places an upper bound on $C$ in the range of 4–17, in agreement with our results.

The experimental results are also confirmed by a microscopic geometric model[24,25] that yields $C = 135\pi^2/64 \approx 20$ if it is assumed that hydraulic conductivity is limited by flow along the faces of cubic solid grains. The transition between cubic and quadratic porosity dependence is dependent on the microscopic melt distribution and, therefore, on chemistry[3], but for the anhydrous olivine/basalt system full connectivity is achieved for porosities of 2–3%. This is reported in ref. 5, which posits that the variation in the nature of the microscopic hydraulic bonds with porosity may lead to a large increase in permeability when full connectivity is achieved. The relatively high porosities of our experiments preclude observation of this transition, but as the grain boundaries in our experiments are saturated in melt, we conclude that $C$ is unlikely to be much lower than the value inferred from our experiments.

Taken at face value, our preferred value of $C = 10$ results in sub-MOR melt transport speeds of $2–150\,\mathrm{m\,yr^{-1}}$ for uniform flow, assuming plausible upper-asthenospheric parameters ($d = 10^{-2}\,\mathrm{m}$ (ref. 26), $\mu = 10\,\mathrm{Pa\,s}$, $\Delta\rho = 500\,\mathrm{kg\,m^{-3}}$) and porosities of $10^{-2\pm0.5}$, as indicated by sub-MOR geophysical observations[8].

**Table 1 | Experimental conditions, calculated melt segregation velocities and permeabilities**

| Run | $T$(°C) | $P$(GPa) | $|\mathbf{a}|/g$ | $t_{static}$ (min) | $t_{centrif}$ (min) | $\phi_0$ | $\dot\phi$ (s⁻¹) | $|\mathbf{v_0}|$ (m s⁻¹) | $k$ (m²) | $d$ (μm) | $C$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| OFB-5 | 1,285 | 1.0 | 1 | 4,320 | — | 0.112(9) | — | — | — | 32.9(10) | — |
| ZOFB-1 | 1,270 | 1.4 | 700 | 1,320 | 273 | 0.051(9) | $3.9 \times 10^{-5}$ | $1.4(6) \times 10^{-8}$ | $2.0(1.1) \times 10^{-14}$ | 22.2(11) | 3.2(1.9) |
| ZOFB-3 | 1,285 | 0.9 | 600 | 1,440 | 251 | 0.097(8) | $3.3 \times 10^{-5}$ | $1.1(3) \times 10^{-8}$ | $2.9(1.2) \times 10^{-14}$ | 21.0(8) | 13.6(5.6) |
| ZOFB-2 | 1,300 | 1.3 | 400 | 1,260 | 627 | 0.121(7) | $9.4 \times 10^{-5}$ | $3.3(9) \times 10^{-9}$ | $2.0(0.8) \times 10^{-14}$ | 17.7(8) | 27(12) |

Here $T$ and $P$ are respectively the experimental pressures and temperatures, $|\mathbf{a}|/g$ is the centrifugal acceleration divided by the acceleration due to gravity at the Earth's surface, $t_{static}$ and $t_{centrif.}$ are respectively the annealing and centrifuging times, $\phi_0$ is the average melt fraction or porosity, $\dot\phi$ is the maximum time-averaged compaction rate (Methods; approximated by $\Delta\phi/t\phi_0$), $\mathbf{v_0}$ is the melt velocity (equation (1)), $k$ is the permeability (equation (2)), $d$ is the time-integrated average grain size during centrifugation (assuming that grain growth is proportional to log $d$/log $t$; this correction amounts to $\sim -5\%$ of the measured grain size at the end of the experiments), and $C$ is the constant in the permeability–porosity relationship (equation (3)).

MOR basalt U-series isotopic excesses are an important indirect source of information about the details of sub-ridge melt flow. Excesses in the long-lived isotopes [230]Th (half-life, $t_{1/2} = 75,000$ yr) and [231]Pa ($t_{1/2} = 32,800$ yr) can be generated by minute degrees of partial melting (<0.1%) in volatile-rich heterogeneities at the base of the melting region at depths of 60–120 km (ref. 27). However, three possibilities are debated for the origin of excesses in more short-lived isotopes, notably [226]Ra ($t_{1/2} = 1,600$ yr). One hypothesis assumes that the more voluminous melting at depths of <60 km, which produces geophysically observable melt fractions of order $10^{-2}$, do not fractionate U-series elements. Consequently, all excesses have a common origin in the lower part of the melting column. Transport times from the bottom of the melting column to the MOR are then constrained by the [226]Ra half-life[12–14], the observed magnitude of [226]Ra excesses suggesting transport in less than twice the [226]Ra half-life. An alternative explanation draws on the observation that the sub-MOR melt flow is strongly channelized[10,28] and proposes that MOR basalt [226]Ra excess is generated by admixture of shallow interchannel partial melts, thereby circumventing the problem of transporting the [226]Ra excess from depths of 120 km (refs 29,30). A third hypothesis explores the generation of excess [226]Ra through diffusive exchanges between melt and minerals in both the asthenospheric melt column and the lithosphere[31]. Our measurements have little bearing on the contribution of these exchange processes to the observed [226]Ra excesses; but because permeability is closely related to transport speed, our measurements constrain the feasibility of a common deep origin for MOR basalt U-series isotopic excesses.

There is consensus that some degree of channelization is necessary to preserve lower-asthenospheric isotopic signatures[30]. Evidence from exhumed asthenospheric rocks[28,32] suggests that this channelization occurs as a consequence of porous-media flow instabilities. Models proposed to explain such instabilities include reactive transport[10,19], shear-enhanced segregation[20] and rheologically induced mechanical instability[9]. We consider the third of these (Methods) because the instability growth rate can be expressed analytically and is independent of external forcings. By this mechanism[9], a domain of anomalous melting initiates a channelling instability. To a first approximation, the growth rate and speed of the instability are respectively proportional to $C^{0.5}$ and $C^{-1}$. The onset of melting would be characterized by low melt viscosities, of 0.1–0.5 Pa s, and background melt fractions of $\phi_0 \approx 10^{-3}$. Taking $C = 10$, as suggested by our analysis, an instability with an initial amplitude of $5\phi_0$ would rise through the lower 60 km of the asthenosphere in 400–1,800 yr, increasing its melt volume about twofold (Fig. 3a). In contrast, for $C = 250$, a value from the previously accepted range for $C$, melt volumes increase three- to five-fold but transport times from depths of 120 to 60 km become 8,000–34,000 yr.

Assuming that such instabilities survive the transition from the lower to the upper asthenosphere, where melt production is more homogeneous, the resulting melts are more viscous ($\mu \approx 10$ Pa s) and background melt fractions are of the order of $\phi_0 \approx 10^{-2}$ (ref. 8), transport times are ~700 yr for $C = 10$, as opposed to ~10,000 yr for $C = 250$ (Fig. 3b). Our results thus show that porous-media channelling instabilities are able to transfer melt from the bottom of the melting column to the MOR in 1,000–2,500 yr, which is fast enough to allow [226]Ra excesses to originate from the bottom of the melting column. Because the velocity scale for compaction-driven

melt transport is directly proportional to the permeability constant, $C$ (ref. 2), we expect that our results also imply an order-of-magnitude decrease in transport timescales inferred for reactive transport[10] and shear-induced melt segregation[20].
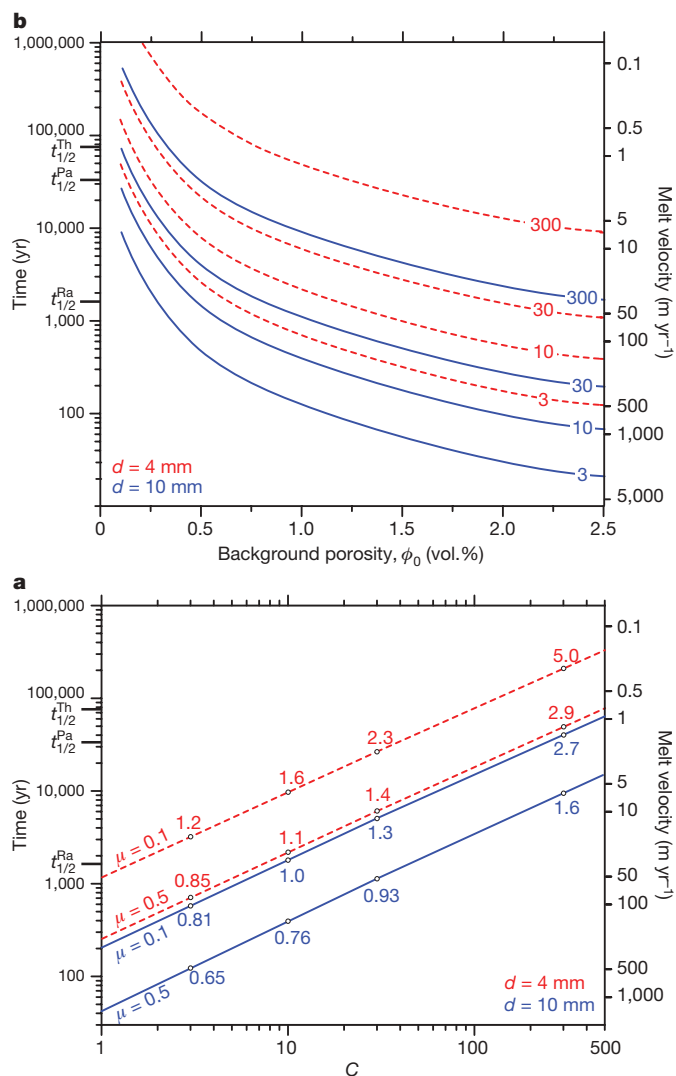


**Figure 3 | Transport times and melt speeds below MORs. a**, Time for a melt channelling instability to traverse the lower MOR asthenosphere (depth, 120–60 km), with background porosity $\phi_0 = 10^{-3}$, as a function of the permeability constant, $C$, grain size, $d$, and melt viscosity, $\mu$. The numbers on individual points give resulting porosities within the instability (in vol.%) at a depth of 60 km. **b**, Time for a melt channelling instability to traverse the upper MOR asthenosphere (depth, 60–0 km) as a function of background porosity, $\phi_0$, grain size, $d$, and the permeability constant, $C$ (labelling the lines). Our preferred value is $C = 10$. The half-lives of [226]Ra, [231]Pa and [230]Th are $t_{1/2}{}^{Ra}$, $t_{1/2}{}^{Pa}$ and $t_{1/2}{}^{Th}$, respectively. In all cases, the instability nucleates from a perturbation with amplitude $5\phi_0$, the solid viscosity is $10^{19}$ Pa s, the difference between the solid and the melt density is 500 kg m⁻³, and the upper-asthenospheric melt viscosity is 10 Pa s. The calculation of transport times is discussed in Methods.

## METHODS SUMMARY

Starting materials were a mixture of natural San Carlos olivine and synthetic MOR basalt glass with a magnesium number (Mg# = $100Mg/(Mg+Fe)$) of 0.80 (Supplementary Table 1). Olivine powder with an average grain size of $\sim$1.8 μm, measured by laser diffractometry of dispersed solutions, was obtained by crushing and milling. The MOR basalt glass was made from natural and synthetic oxide powders in a platinum crucible at 1,300 °C with an oxygen fugacity ($f_{O_2}$) such that $\log f_{O_2} = -6.704$ (equivalent to quartz–fayalite–magnetite at 1,300 °C and 1 GPa).

Experiments for annealing were performed in an end-loaded, 14-mm-bore piston–cylinder, using double platinum–graphite capsules. The bore sizes and salt/pyrex/graphite/crushable MgO assemblies of the static and centrifuging piston–cylinder devices were identical[21], the graphite furnaces were 6.0 mm in inner diameter and 36 mm in length, and the platinum capsules were 4 mm in outer diameter and typically 7 mm in length. The capsules were placed such that the hot spot was centred on the sample. Temperature was controlled to a precision of $\pm$5 °C and the temperature gradient along the length of the capsule was <10 °C. The arrangement in the centrifuging piston–cylinder was such that the assembly axis was radial, with the thermocouple towards the centre, the bottom of the capsule facing the thermocouple and acceleration directed along the axis of the capsule towards its lid.

**Melt distribution.** Polished sections of the run products were imaged using backscattered electrons and olivines were analysed by wavelength-dispersive spectrometry using a Jeol JXA-8200 microprobe. Melt pools were too small to analyse. Melt abundances along the capsule and average grain sizes of olivines were obtained through digital analysis of backscattered-electron images using the software IMAGETOOL version 3.0 (The University of Texas Health Science Center; http://ddsdx.uthscsa.edu/dig/itdesc.html).

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Maaloe, S. & Scheie, A. The permeability controlled accumulation of primary magma. *Contrib. Mineral. Petrol.* **81**, 350–357 (1982).
2. McKenzie, D. The generation and compaction of partially molten rock. *J. Petrol.* **25**, 713–765 (1984).
3. Van Bargen, N. & Waff, H. S. Permeabilities, interfacial areas and curvatures of partially molten systems: results of numerical computations of equilibrium microstructures. *J. Geophys. Res.* **91**, 9261–9276 (1986).
4. Wark, D. A., Williams, C. A., Watson, E. B. & Price, J. D. Reassessment of pore shapes in microstructurally equilibrated rocks, with implications for permeability of the upper mantle. *J. Geophys. Res.* **108**, doi:10.1029/2001JB001575 (2003).
5. Faul, U. H. Melt retention and segregation beneath mid-ocean ridges. *Nature* **410**, 920–923 (2001).
6. Cheadle, M. J., Elliott, M. T. & McKenzie, D. Percolation threshold and permeability of crystallizing igneous rocks: the importance of textural equilibrium. *Geology* **32**, 757–760 (2004).
7. Richardson, C. & McKenzie, D. Radioactive disequilibria from 2D models of melt generation by plumes and ridges. *Earth Planet. Sci. Lett.* **128**, 425–437 (1994).
8. The MELT Seismic Team. Imaging the deep seismic structure beneath a mid-ocean ridge: the MELT experiment. *Science* **280**, 1215–1218 (1998).
9. Connolly, J. A. D. & Podladchikov, Y. Y. Decompaction weakening and channeling instability in ductile porous media: implications for asthenospheric melt segregation. *J. Geophys. Res.* **112**, doi:10.1029/2005JB004213 (2007).
10. Kelemen, P. B., Hirth, G., Shimizu, N., Spiegelman, M. & Dick, H. J. B. A review of melt migration processes in the adiabatically upwelling mantle beneath oceanic spreading ridges. *Phil. Trans. R. Soc. Lond. A* **355**, 283–318 (1997).
11. Spiegelman, M., Kelemen, P. B. & Aharonov, E. Causes and consequences of flow organization during melt transport: the reaction infiltration instability in compactible media. *J. Geophys. Res.* **106**, 2061–2077 (2001).
12. McKenzie, D. Constraints on melt generation and transport from U-series activity ratios. *Chem. Geol.* **162**, 81–94 (2000).
13. Rubin, K. H., van der Zander, I., Smith, M. C. & Bergmanis, E. C. Minimum speed limit for ocean ridge magmatism from $^{210}$Pb–$^{226}$Ra–$^{230}$Th disequilibria. *Nature* **437**, 534–538 (2005).
14. Stracke, A., Bourdon, B. & McKenzie, D. Melt extraction in the Earth's mantle: constraints from U–Th–Pa–Ra studies in oceanic basalts. *Earth Planet. Sci. Lett.* **244**, 97–112 (2006).
15. Spiegelman, M. & McKenzie, D. Simple 2-D models for melt extraction at mid-ocean ridges and island arcs. *Earth Planet. Sci. Lett.* **83**, 136–152 (1987).
16. Johnson, K. T. M., Dick, H. J. B. & Shimizu, N. Melting in the oceanic upper mantle: an ion microprobe study of diopsides in abyssal peridotites. *J. Geophys. Res.* **95**, 2661–2678 (1990).
17. Sobolev, A. V. & Shimizu, N. Ultra-depleted primary melt included in an olivine from the Mid-Atlantic Ridge. *Nature* **363**, 151–154 (1999).
18. Salters, V. J. M. & Longhi, J. Trace element partitioning during the initial stages of melting beneath mid-ocean ridges. *Earth Planet. Sci. Lett.* **166**, 15–30 (1999).
19. Daines, M. J. & Kohlstedt, D. L. The transition from porous to channelized flow due to melt/rock reaction during melt migration. *Geophys. Res. Lett.* **21**, 145–148 (1994).
20. Holtzman, B. K. Groebner, N. J., Zimmermann, M. E., Ginsberg, S. B. & Kohlstedt, D. L. Stress-driven melt segregation in partially molten rocks. *Geochem. Geophys. Geosyst.* **4**, doi:10.1029/2001GC000258 (2003).
21. Schmidt, M. W., Connolly, J. A. D., Günter, D. & Bogaerts, M. Element partitioning: the role of melt structure and composition. *Science* **312**, 1646–1650 (2006).
22. Renner, J. Viskupic, K. Hirth, G. & Evans, B. Melt extraction from partially molten peridotites. *Geochem. Geophys. Geosyst.* **4**, doi:10.1029/2002GC000369 (2003).
23. Bottinga, Y. & Weill, D. F. The viscosity of magmatic silicate liquids: a model for calculation. *Am. J. Sci.* **272**, 438–475 (1972).
24. Dienes, J. K. in *Issues in Rock Mechanics* (eds Goodman, R. E. & Heuze, F. E.) 86–94 (American Institute of Mining, Metallurgical, and Petroleum Engineers, 1982).
25. Connolly, J. A. D., Holness, M. B., Rubie, D. C. & Rushmer, T. Reaction-induced microcracking: an experimental investigation of a mechanism for enhancing anatectic melt extraction. *Geology* **25**, 591–594 (1997).
26. Mercier, J. C. Magnitude of the continental lithospheric stresses inferred from rheomorphic petrology. *J. Geophys. Res.* **85**, 6293–6303 (1980).
27. Hirth, G. & Kohlstedt, D. L. Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere. *Earth Planet. Sci. Lett.* **144**, 93–108 (1996).
28. Kelemen, P. B., Shimizu, N. & Salters, V. J. M. Extraction of mid-ocean-ridge basalt from the upwelling mantle by focused flow of melt in dunite channels. *Nature* **375**, 747–753 (1995).
29. Spiegelman, M. & Elliott, T. Consequences of melt transport for uranium series disequilibrium in young lavas. *Earth Planet. Sci. Lett.* **118**, 1–20 (1993).
30. Lundstrom, C. C. Uranium-series disequilibria in mid-ocean ridge basalts: observations and models of basalt genesis. *Rev. Mineral. Geochem.* **52**, 175–214 (2003).
31. Van Orman, J. A., Saal, A. E., Bourdon, B. & Hauri, E. H. Diffusive fractionation of U-series radionuclides during mantle melting and shallow-level melt-cumulate interaction. *Geochim. Cosmochim. Acta* **70**, 4797–4812 (2006).
32. Bouilhol, P. et al. Recording of arc crust-mantle transition zone formation by melt-rock reaction: evidence from ultramafic rocks of Sapat (Kohistan, Northern Pakistan). *Lithos* **107**, 17–37 (2009).

## METHODS

**Fluid velocity relative to the sample barycentre.** Equation (1) is obtained by integrating the conservation equation for melt mass

$$\frac{\partial \phi}{\partial t} + \nabla \bullet (\phi \mathbf{v}) = 0 \qquad (4)$$

over time and space. Equation (1) is the melt velocity relative to the barycentre rather than the true Darcyian velocity, $\mathbf{v}$, which is measured relative to the solid matrix. However, for one-dimensional compaction involving incompressible phases, conservation of total mass relates the solid and fluid velocities at the volumetric barycentre as follows:

$$(1-\phi_0)\mathbf{v}_{0,s} + \mathbf{v}_0\phi_0 = 0 \qquad (5)$$

where $\mathbf{v}_{0,s}$ is the barycentric solid velocity. Thus, assuming the maximum possible hydraulic gradient of $-(1-\phi)\Delta\rho\mathbf{a}$ for a self-compacting system, Darcy's law is

$$(\mathbf{v}_0 - \mathbf{v}_{0,s})\phi_0 = \frac{k}{\mu}(1-\phi_0)\Delta\rho\mathbf{a} \qquad (6)$$

By making use of equation (5) to express the barycentric solid velocity as a function of $\mathbf{v}_0$ and $\phi_0$, equation (6) can be simplified to equation (2).

**Rheologically limited compaction rate.** When the matrix has a much greater viscosity than its interstitial fluid, the viscous compaction rate[33,34] is

$$\dot{\phi} = \frac{1}{\phi}\frac{d\phi}{dt} = -\frac{3}{4}\frac{p_e}{\eta} \qquad (7)$$

where $p_e$ is the effective pressure and $\eta$ is the shear viscosity of the solid grains. We compute the olivine shear viscosity from the constitutive relation

$$\eta = \frac{d^3}{3A}\exp\left(\frac{E+pV}{RT}\right) \qquad (8)$$

for diffusion creep in dry olivine as constrained experimentally[35], where $p$ is pressure (in pascals), $T$ is temperature (in kelvin), $R$ is the molar gas constant, $d$ is grain size (in metres), $E = 375 \times 10^3\,\mathrm{J\,mol^{-1}}$, $V = 6 \times 10^{-6}\,\mathrm{m^3}$ and $A = 1.5 \times 10^{-15}\,\mathrm{m^3\,Pa^{-1}\,s^{-1}}$. For rheologically limited compaction[34], the effective pressure is

$$p_e = -\int(1-\phi)\Delta\rho|\mathbf{a}|\,dz \approx -\Delta\rho|\mathbf{a}|z$$

and attains its greatest magnitude at the top and bottom of the sample, $z = \pm H$. The observed compaction rates (Table 1) are $10^5$ (ZOFB-3) to $10^6$ (ZOFB-1) times faster than those predicted from equations (7) and (8). Although dry olivine is not necessarily an appropriate model for compaction of a melt-saturated matrix; the compaction rates reported in ref. 22 for olivine plus basalt melt are in good agreement with those computed from equations (7) and (8), with steady-state compaction rates being less than two orders of magnitude greater than predicted. This suggests that although viscous creep plausibly explains the observations in ref. 22, compaction must have been accelerated in our experiments by a plastic mechanism activated by centrifuging.

**Mechanical channelling instability.** To quantify the properties of buoyancy-induced channelling instabilities that result from compaction-driven fluid flow through a viscous matrix that weakens under high fluid pressure, we follow ref. 9 in characterizing the weakening by a reduced viscosity during decompaction. For this case, numerical modelling[9] has demonstrated that in porous media with a cubic porosity–permeability relationship, a domain of anomalous melting initiates an instability that grows such that the amplitude of the instability measured relative to the background melt fraction, $\phi_0$, at time $t$ is

$$\alpha \approx \alpha_0 + \frac{t}{\kappa^{3/8}}\sqrt{\frac{C\mu}{d^2\phi_0^2\eta}} \qquad (9)$$

where $\alpha$ is the amplitude of the instability measured relative to the background melt fraction, $\phi_0$, at time $t$, and we set $\eta = 10^{19}\,\mathrm{Pa\,s}$ and $d = 0.01\,\mathrm{m}$. The parameter $\kappa$ characterizes weakening during decompaction. A value of $\kappa < 10^{-2}$–$10^{-3}$ is necessary to reproduce channelization patterns inferred from natural systems[10,32]; accordingly, we take $\kappa = 10^{-3}$. The speed of the instability is

$$c = \frac{d^2\phi_0^2\Delta\rho g}{C\mu}\frac{1 + (3\ln\alpha - 1)\alpha^3}{1 + (\alpha^2 - 3)\alpha/2}$$

The time required for an instability of initial amplitude $\alpha_0$ to propagate a distance $\delta z$ is therefore obtained by solving the transcendental equation

$$\delta z - \int c\,dt = 0 \qquad (10)$$

This analysis presumes that the porosity distribution generated by anomalous melting is close to that of the quasi-steady-state solution observed numerically. When this assumption is invalid, the time required for the porosity distribution to evolve towards a quasi-steady state is of the order of the compaction timescale. As the instability grows by draining melt from the background porosity, short-lived radiogenic isotope excesses (for example for $^{226}$Ra) can be generated after nucleation. The volume of melt transported is approximately linearly proportional to the instability amplitude.

Whether such instabilities would persist in the upper asthenosphere, where melt production is homogeneous, the melts are more viscous and background melt fractions are $\sim 10^{-2}$, remains to be demonstrated. However, assuming that the instabilities survive the transition, the initial amplitude of the instability in the upper regime would be

$$\alpha'_0 \approx \alpha\left(\frac{\phi_0}{\phi'_0}\right)^4\left(\frac{\mu'}{\mu}\right)^{3/2} \qquad (11)$$

where primes denote upper asthenospheric properties.

33. Wilkinson, D. S. & Ashby, M. F. Pressure sintering by power law creep. *Acta Metall.* **23,** 1277–1285 (1975).
34. Connolly, J. A. D. & Podladchikov, Y. Y. Temperature-dependent viscoelastic compaction and compartmentalization in sedimentary basins. *Tectonophysics* **324,** 137–168 (2000).
35. Hirth, G. & Kohlstedt, D. L. in *Inside the Subduction Factory* (ed. Eiler, J. M.) 83–106 (Geophys. Monogr. 138, American Geophysical Union, 2003).

# Human-specific transcriptional regulation of CNS development genes by FOXP2

Genevieve Konopka[1,3], Jamee M. Bomar[1,3], Kellen Winden[1,3], Giovanni Coppola[3], Zophonias O. Jonsson[5], Fuying Gao[3], Sophia Peng[3], Todd M. Preuss[6], James A. Wohlschlegel[5] & Daniel H. Geschwind[1,2,3,4]

The signalling pathways controlling both the evolution and development of language in the human brain remain unknown. So far, the transcription factor *FOXP2* (forkhead box P2) is the only gene implicated in Mendelian forms of human speech and language dysfunction[1–3]. It has been proposed that the amino acid composition in the human variant of FOXP2 has undergone accelerated evolution, and this two-amino-acid change occurred around the time of language emergence in humans[4,5]. However, this remains controversial, and whether the acquisition of these amino acids in human FOXP2 has any functional consequence in human neurons remains untested. Here we demonstrate that these two human-specific amino acids alter FOXP2 function by conferring differential transcriptional regulation *in vitro*. We extend these observations *in vivo* to human and chimpanzee brain, and use network analysis to identify novel relationships among the differentially expressed genes. These data provide experimental support for the functional relevance of changes in FOXP2 that occur on the human lineage, highlighting specific pathways with direct consequences for human brain development and disease in the central nervous system (CNS). Because FOXP2 has an important role in speech and language in humans, the identified targets may have a critical function in the development and evolution of language circuitry in humans.

The amino acid structure of FOXP2 had been highly conserved along the mammalian lineage until the common ancestor of humans and chimpanzees, when the human variant of FOXP2 acquired two different amino acids under positive selection, which has been interpreted as evidence for accelerated evolution[4,5]. To test whether the amino acids under positive selection in human FOXP2 have a distinct biological function, which would support the role of these changes in evolution, we expressed either human FOXP2 or the same construct mutated at two sites to yield the chimpanzee amino acid content, FOXP2[chimp], in human neuronal cells without endogenous FOXP2 (Fig. 1a–f). Exogenous FOXP2 protein expressed from both constructs was localized in the nucleus as determined by immunocytochemistry (Fig. 1c–e) and subcellular fractionation (Fig. 1f), consistent with its endogenous expression. To determine if modifying two amino acids leads to changes in gene expression, we conducted whole-genome microarray analysis. We identified 61 genes significantly upregulated and 55 genes downregulated by FOXP2 compared to FOXP2[chimp] (Supplementary Table 1), as well as genes regulated by both FOXP2 and FOXP2[chimp] (Supplementary Table 2). Notably, FOXP2[chimp] overexpression resulted in more changes in gene regulation than FOXP2 (Supplementary Table 3). In replicate experiments in a different human neuronal cell line, FOXP2[chimp] again regulated more genes than FOXP2 even though its expression was

higher than FOXP2 in these cells (data not shown). To control for any potential confounding effects of FOXP2 levels, we performed correlations of the levels of every gene on the array to either FOXP2 or FOXP2[chimp] levels, as well as performing random permutation testing, and found no significant differences between other genes' correlations to either FOXP2 or FOXP2[chimp]. These data indicate that the differentially expressed genes are not due to different levels of FOXP2 or FOXP2[chimp], and are a true indication of differential transcriptional regulation by these two proteins.

To confirm the validity of differentially expressed FOXP2 target genes, we conducted quantitative polymerase chain reaction following reverse transcription (qRT–PCR) using independent RNA samples. We confirmed 93% of the FOXP2 upregulated genes and 75% of the downregulated genes examined (Fig. 1g, h and Supplementary Fig. 1). Five genes confirmed by qRT–PCR (*COL9A1*, *ROR2*, *SLIT1*, *SYK*, and *TAGLN*; Fig. 1g, h and Supplementary Fig. 1) were previously identified as direct FOXP2 targets using ChIP-chip[6,7]. Sixty per cent of promoters of the identified differentially expressed genes have at least one canonical FOXP2 binding site, 92% have at least one forkhead domain binding site, and 99% have at least one 'core' FOXP2 binding site (Supplementary Table 4). The canonical FOXP2 binding site CAAATT, as well as the core site AAAT, is significantly enriched in the downregulated genes ($P = 3.3 \times 10^{-4}$ and $P = 8.6 \times 10^{-3}$, respectively) compared to randomly permuting the same number of promoters from the genome. Genes with promoters containing a canonical FOXP2 binding site are likely to be direct FOXP2 or FOXP2[chimp] targets.

To confirm that these findings were not an artefact of the cell lines used, we further assessed whether a different primary neural cell, human neural progenitors, would show similar differential regulation by FOXP2 and FOXP2[chimp]. We confirmed one-third of the genes examined in these human cells using both a different method of gene transduction, and populations of cells with greater levels of FOXP2[chimp] compared to human FOXP2 overexpression, which complements the SH-SY5Y data to further show that the observed relationships are not due to FOXP2 levels (Supplementary Fig. 2). As an additional level of validation and to extend the findings to the level of protein, we confirmed two genes, *CACNB2* and *ENPP2*, by immunoblotting in additional SH-SY5Y cell lines (Supplementary Fig. 3).

To explore the potential function of the differential FOXP2 targets, we determined enrichment of gene ontology (GO) categories. GO categories enriched for genes upregulated by FOXP2 compared to FOXP2[chimp] are involved in transcriptional regulation of gene expression and cell–cell signalling. Those GO categories enriched for genes downregulated by FOXP2 compared to FOXP2[chimp] are important for protein and cell regulation (Supplementary Table 5).

[1]Program in Neurogenetics, [2]Semel Institute and Department of Psychiatry, [3]Departments of Neurology, [4]Human Genetics, and [5]Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. [6]Division of Neuroscience and Center for Behavioral Neuroscience, Yerkes National Primate Research Center, and Department of Pathology & Laboratory Medicine, Emory University School of Medicine, Atlanta, Georgia 30329, USA.
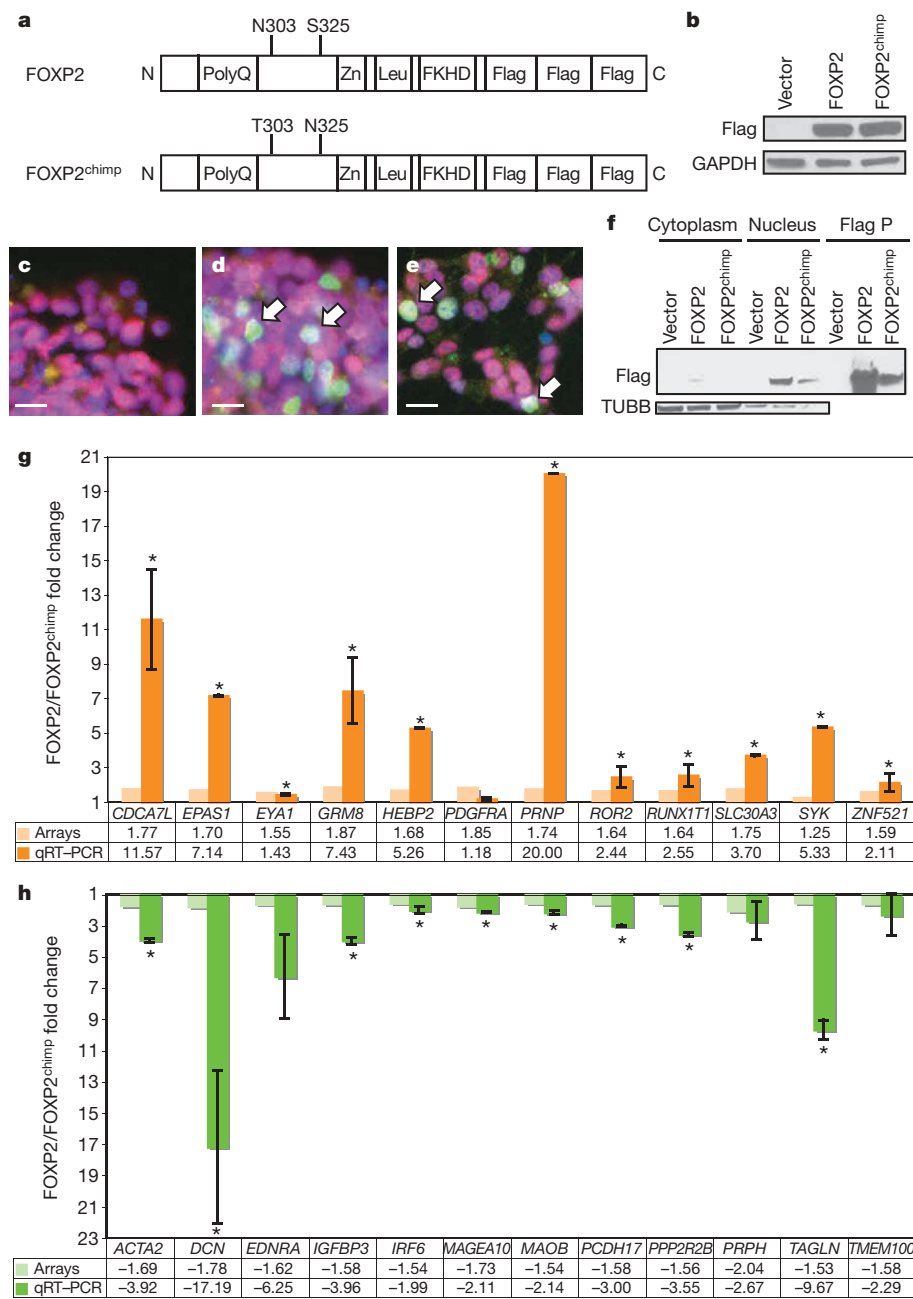
**Figure 1 | FOXP2 and FOXP2chimp differentially regulate genes in SH-SY5Y cells. a**, Schematic of human FOXP2 showing its major functional protein domains (Zn indicates the zinc finger domain, Leu indicates the leucine-zipper domain, and FKHD indicates the forkhead DNA binding domain) and the two amino acid changes in the mutant FOXP2chimp. **b**, Representative immunoblot for Flag-tagged FOXP2 and FOXP2chimp stable overexpression in SH-SY5Y cells. **c–e**, Immunofluorescent staining of antibodies against Flag epitope (green) and FOXP1 (red), and 4,6-diamidino-2-phenylindole (DAPI; blue) for nuclei. Vector cells demonstrate no Flag expression (**c**), whereas both FOXP2 (**d**) and FOXP2chimp (**e**) expressing cells have Flag-tagged FOXP2 in the cell nucleus. Arrows indicate examples of cell nuclei positive for Flag expression. Scale bars, 5 μm. **f**, Subcellular fractionation followed by immunoblotting. **g, h**, Quantitative RT–PCR of genes that were differentially expressed in cells expressing FOXP2 compared to FOXP2chimp. Asterisks indicate $P \leq 0.05$ and error bars are ± s.e.m. (two-tailed Student's $t$-test, $n = 3$ or 4).

These data support the idea that FOXP2 and FOXP2chimp have distinguishable downstream effects as reflected by their differences in gene regulation.

To determine the potential mechanisms by which FOXP2 or FOXP2chimp might differentially regulate gene expression, we first examined whether either protein preferentially interacts with FOXP1 or FOXP4, two proteins known to form a heterodimer with FOXP2 (ref. 8). Both FOXP2 and FOXP2chimp co-localized with FOXP1 in the cell nucleus, co-immunoprecipitated with FOXP1 as evidenced by immunoblotting, and co-immunoprecipitated with both FOXP1 and FOXP4 when assayed by mass spectrometry (Figs 1c–e, 2a, b and Supplementary Fig. 4b–g), ruling out a major

difference in FOXP1 or FOXP4 interaction causing differential gene expression. Mass spectrometry showed no significant difference in either co-immunoprecipitation experiment, indicating that differences in hetero- or homodimerization did not underlie the observed differences in gene expression between the chimpanzee and human FOXP2. We also tested whether changes in cell proliferation could account for gene expression differences, but did not find significant changes in growth with either FOXP2 construct (Fig. 2c).

We next assessed whether FOXP2 and FOXP2chimp expression led to differential promoter transactivation of target genes. We selected eight genes confirmed by qRT–PCR that also contained at least one forkhead binding site (Supplementary Table 6). Six of the promoters
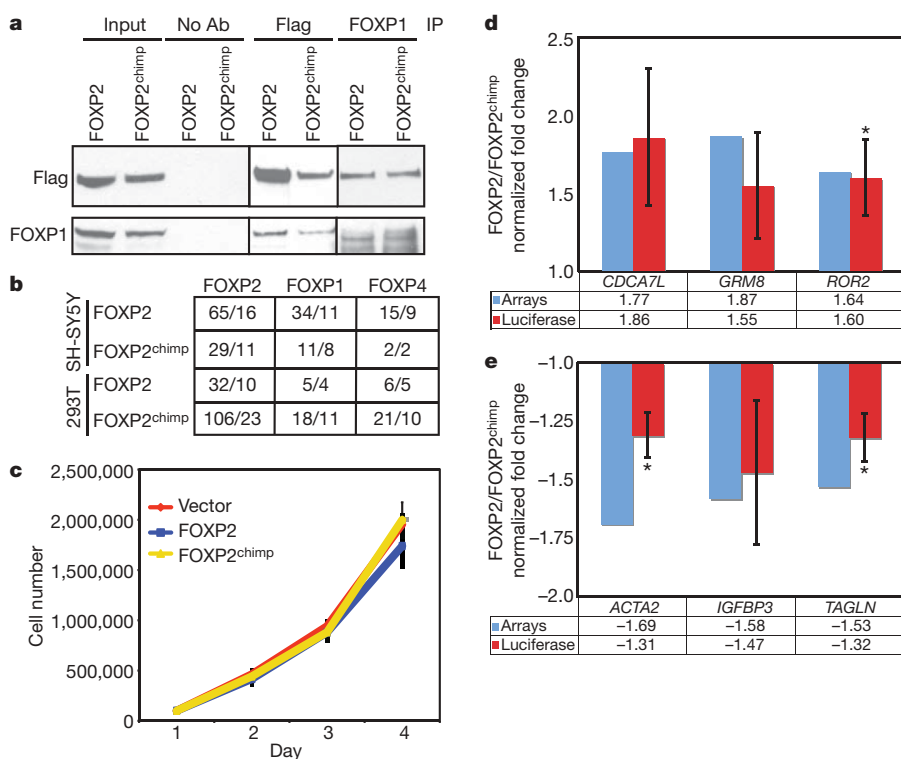
**Figure 2 | FOXP2 and FOXP2<sup>chimp</sup> differentially transactivate target promoters independent of FOXP1 or FOXP4 interaction. a**, Immunoblotting for Flag or FOXP1 following immunoprecipitation with either Flag or FOXP1 antibodies. **b**, Mass spectrometry results from SH-SY5Y or 293T cells overexpressing FOXP2 or FOXP2<sup>chimp</sup>. The first number indicates the number of spectra and the second is the number of unique peptides. **c**, Cell growth analysis does not show a significant difference in proliferation between cells expressing FOXP2 or FOXP2<sup>chimp</sup> over time ($P \leq 0.05$). Error bars are $\pm$ s.e.m. (two-tailed Student's $t$-test, $n = 3$). **d, e**, Dual luciferase assays in 293T cells transiently transfected with promoter fragments driving luciferase and either FOXP2 or FOXP2<sup>chimp</sup>. Asterisks indicate $P \leq 0.05$ and error bars are $\pm$ s.e.m. (two-tailed Student's $t$-test, $n = 3–6$).

tested showed differential regulation by FOXP2 compared to FOXP2[chimp] in the same direction as the microarrays (Fig. 2d, e), whereas two did not demonstrate significant transactivation in either direction (data not shown). In contrast, a canonical FOXP2 binding site in triplicate alone, outside of a genomic context, was regulated equally by both FOXP2 and FOXP2[chimp] (Supplementary Fig. 5). Given the complexity of *cis*-acting gene transactivation elements, these data are particularly compelling considering our use of simplified 5′ promoter regions. These data demonstrate that at least a subset of differentially regulated genes is also differentially transactivated by FOXP2 and FOXP2[chimp], indicating that they are probably direct FOXP2 targets.

To place these gene expression changes within a more systematic context, we applied weighted gene co-expression network analysis[9,10] to the entire SH-SY5Y microarray data set to examine co-regulation of gene expression across all genes. We uncovered two modules where the module eigengene (for definition, see Methods) was driven by differences in FOXP2 and FOXP2[chimp], and one module driven by similar gene regulation (Fig. 3 and Supplementary Fig. 6). Using this unsupervised analysis, we found additional genes of interest that do not meet the criteria for differential expression, but that are co-regulated with differences in FOXP2 and FOXP2[chimp] expression (Supplementary Table 7). Notably, two of the genes with the most connections, so-called 'hub' genes, in one of the differential networks are *DLX5* and *SYT4*, two genes important for brain development and function[11,12].

To extrapolate these findings to true *in vivo* expression and provide external validation, we compared the differentially expressed genes in SH-SY5Y cells to differentially expressed genes from adult human and chimpanzee brain tissue. We performed microarray analysis on tissue from three brain regions where FOXP2 is expressed in developing brain: caudate nucleus, frontal pole and hippocampus. We examined gene expression in human compared to chimpanzee for each brain region separately as well as for all brain regions combined, for a total of eight comparisons. There was a significant overlap in seven out of eight of these comparisons, a remarkable convergence with the *in vitro* data (Table 1). These data are particularly notable, as the tissue was from adult brain. We surmise that a subset of the overlapping differentially expressed genes found in adult brain is the result of differential functions by FOXP2 in the developing brain,

and may lead to increased vulnerability to disease. For example, mutations in both *FGF14* and *PPP2R2B* lead to spinocerebellar ataxia (spinocerebellar ataxia type 27 and 12, respectively), which involves motor-related speech defects[13,14]. Because both of these genes have a critical role in cerebellar function, it is of note that patients with FOXP2 mutations have decreased grey matter in the cerebellum[15], and *Foxp2* knockout mice have their most pronounced morphological phenotype in the cerebellum[16]. Mutations in *COL9A1* lead to Stickler syndrome in which patients have craniofacial abnormalities[17], and patients with mutations in *GJA12* (also called *GJC2*) present with ataxia, nystagmus, other motor impairments, and often mental retardation[18].

Although comparisons of developing brain between human and chimpanzee are challenged by a lack of tissue, a recent study examined gene expression in many regions of human fetal brain[19]. Comparing the list of 116 differentially expressed genes with those focally expressed during human fetal development, we find 14 genes specifically expressed in one brain region, including FOXP2 (Supplementary Table 8). Two regions of the human fetal brain with high FOXP2 expression[19]—perisylvian cortex and cerebellum—have a significant number of enriched genes that overlap with the differentially expressed FOXP2 and FOXP2[chimp] genes ($P = 1.1 \times 10^{-4}$ and $P = 1.3 \times 10^{-4}$, respectively; Supplementary Table 8). A significant number of the differentially expressed genes are also associated with human-specific accelerated highly conserved non-coding sequences (haCNS), but not with chimpanzee highly conserved non-coding sequences ($P = 1.2 \times 10^{-6}$ and $P = 0.04$; Supplementary Table 8)[19,20]. We confirmed a number of these genes, such as *GRM8*, *MAOB*, *PPP2R2B*, *PRICKLE1* and *RUNX1T1*, either by qRT–PCR and/or with the adult *in vivo* data set (Fig. 1 and Table 1). Together, these data suggest that the FOXP2 differentially expressed genes identified here may have important roles in brain development and patterning, and may also have evolved *cis*-regulatory elements important for their expression specifically in human brain.

Previously, we identified ChIP-chip targets of FOXP2 that themselves were also under positive selection[6]. We hypothesized that networks of genes important for language circuitry had been positively selected through selective pressure on human brain evolution. Thus, we also examined whether any differential FOXP2 targets were themselves

**Figure 3 | Visualization of one of the modules containing FOXP2 and FOXP2chimp differentially expressed genes.** Two-hundred-and-fifty pairs of genes with the greatest topological overlap are shown. Positive correlations are depicted in red and negative correlations are depicted in blue. The gene symbols for hub genes are accentuated in large, bold text.

under positive selection. Five genes (*AMT*, *C6orf48*, *MAGEA10*, *PHACTR2* and *SH3PXD2B*) met the standard criteria of $K_a/K_s \geq 1.0$ for positive selection on the human lineage (where $K_a$ indicates the rate of non-synonymous substitutions and $K_s$ indicates the rate of synonymous substitutions; Supplementary Table 9)[21]. These data, along with the haCNS and expression data mentioned above, suggest that a subset of differential FOXP2 targets may have co-evolved to regulate pathways involved in higher cognitive functions.

The positive selection of two amino acids in human FOXP2 was previously hypothesized as a mechanism by which human FOXP2 might assume a novel biological function with implications for speech and language evolution[4,5]. A recent study made an elegant attempt to examine the role of these two amino acids by generating a transgenic mouse with the human version of FOXP2 (ref. 22). These mice show a number of interesting phenotypic alterations including increases in dendritic length in striatal neurons and changes in

**Table 1 | Overlap of cell and *in vivo* microarray data**

| Genes | Cells | All brain areas (7.36 × 10⁻⁴ upregulated; 1.10 × 10⁻⁶ downregulated)* | Hippocampus (4.49 × 10⁻² upregulated; 4.07 × 10⁻⁸ downregulated)* | Caudate (4.04 × 10⁻² upregulated; 2.86 × 10⁻² downregulated)* | Frontal pole (7.21 × 10⁻² upregulated; 1.86 × 10⁻⁴ downregulated)* |
|---|---|---|---|---|---|
| **Upregulated** | | | | | |
| ADAMTS9 | 1.38 | 1.91 | 1.72 | 2.35 | 1.73 |
| BCAN | 1.29 | – | – | 1.64 | – |
| COL9A1 | 1.26 | 1.23 | – | 1.26 | – |
| EXPH5 | 1.27 | 1.41 | 1.29 | 1.43 | 1.53 |
| FRZB | 1.32 | 1.65 | 1.39 | 2.04 | 1.57 |
| IGFBP4 | 1.27 | 1.35 | – | – | 1.59 |
| ISLR2 | 1.30 | 1.27 | 1.53 | – | – |
| MGST1 | 1.26 | 2.33 | 1.79 | 3.34 | 2.13 |
| NPTX2 | 1.24 | 1.28 | – | 1.46 | – |
| PDGFRA | 1.84 | 1.27 | – | 1.35 | – |
| PRICKLE1 | 1.45 | – | 1.43 | – | – |
| RUNX1T1 | 1.64 | 1.24 | 1.33 | – | 1.24 |
| SLC30A3 | 1.75 | 1.92 | 2.20 | 1.83 | 1.74 |
| **Downregulated** | | | | | |
| ACCN2 | −1.27 | −1.33 | −1.31 | −1.28 | −1.39 |
| B3GNT1 | −1.24 | −1.73 | −1.47 | −2.55 | −1.39 |
| C6orf48 | −1.23 | −1.56 | −1.53 | −1.65 | −1.51 |
| C8orf13 | −1.27 | −1.36 | −1.64 | – | −1.29 |
| CACNB2 | −1.23 | −1.69 | −1.4 | −2.58 | −1.35 |
| DCN | −1.78 | −1.39 | −1.68 | – | −1.33 |
| ELMO1 | −1.28 | −1.34 | – | −1.64 | −1.32 |
| ENPP2 | −1.39 | −1.43 | −1.72 | – | – |
| FAM43A | −1.40 | – | −1.32 | – | – |
| FAM43B | −1.40 | – | – | – | −1.41 |
| FGF14 | −1.23 | −1.34 | −1.24 | −1.57 | – |
| FLJ11286 | −1.27 | −1.38 | −1.32 | −1.31 | −1.51 |
| GJA12 | −1.25 | −1.35 | – | – | – |
| GLRX | −1.29 | – | – | −1.30 | – |
| HIST2H2BE | −1.30 | −1.39 | −1.61 | – | −1.39 |
| IFIT2 | −1.24 | −1.34 | −1.49 | – | – |
| IGFBP3 | −1.58 | – | −1.30 | – | – |
| MAOB | −1.54 | – | −1.26 | – | – |
| PPP2R2B | −1.56 | −1.23 | −1.66 | −1.35 | −1.41 |

* Overlap *P* values for upregulated and downregulated genes are given in parentheses.

ultrasonic vocalizations, as well as some modest changes in gene expression. Although the mouse is an experimentally tractable model system, from a strictly evolutionary standpoint, the interpretation of data obtained in the mouse specifically for the study of human evolution is challenged by the vast differences in human and mouse brain and the amount of time since the human and mouse common ancestor diverged (70 million years[23]). Here, we demonstrate that these two amino acid changes have a functional consequence in human cells, validate these differences *in vivo* in tissue, and elucidate some of the downstream pathways affected by this adaptive evolutionary change.

Using whole-genome microarrays, we uncovered genes that are differentially regulated upon mutation of these two amino acids, including some with functions critical to the development of the human central nervous system. Moreover, this study reveals enrichment of differential FOXP2 targets with known involvement in cerebellar motor function, craniofacial formation, and cartilage and connective tissue formation, suggesting an important role for human FOXP2 in establishing both the neural circuitry and physical structures needed for spoken language. The significant overlap of human FOXP2 targets in cell lines with genes enriched in human compared to chimpanzee brain tissue presents the possibility that human and chimpanzee FOXP2 have differentially regulated targets during brain development. As suggested over 30 years ago[24], and reaffirmed by the sequencing of both the human and chimpanzee genomes, the phenotypic differences exhibited by humans and chimpanzees cannot be explained by differences in DNA sequence alone, and are probably due to differences in gene expression and regulation. Previous microarray studies identified differences in gene expression between human and chimpanzee brains[25,26]. Here, we link new whole-genome expression microarray data from human and chimpanzee brain to direct differences in gene regulation by the human and chimpanzee version of the transcription factor FOXP2. Because normal FOXP2 function is critical for speech in humans, these differentially regulated targets may be relevant to the evolution and establishment or function of pathways necessary for speech and language in humans.

## METHODS SUMMARY

**Cell culture and stable line generation.** SH-SY5Y cells (ATCC) and human fetal neuronal progenitors (Lonza) were grown according to the manufacturer's instructions, with some modifications (see Methods).

**Microarrays.** Total RNA was extracted using Qiagen's RNeasy kit. Illumina HumanRef-8 v2 (SH-SY5Y samples) or v3 (tissue samples) were used and analysed as described[27]. Sample information is in Methods.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Feuk, L. *et al.* Absence of a paternally inherited FOXP2 gene in developmental verbal dyspraxia. *Am. J. Hum. Genet.* 79, 965–972 (2006).
2. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523 (2001).
3. MacDermot, K. D. *et al.* Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.* 76, 1074–1080 (2005).
4. Enard, W. *et al.* Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418, 869–872 (2002).
5. Zhang, J., Webb, D. M. & Podlaha, O. Accelerated protein evolution and origins of human-specific features: Foxp2 as an example. *Genetics* 162, 1825–1835 (2002).
6. Spiteri, E. *et al.* Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am. J. Hum. Genet.* 81, 1144–1157 (2007).
7. Vernes, S. C. *et al.* High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *Am. J. Hum. Genet.* 81, 1232–1250 (2007).
8. Li, S., Weidenfeld, J. & Morrisey, E. E. Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. *Mol. Cell. Biol.* 24, 809–822 (2004).
9. Oldham, M. C. *et al.* Functional organization of the transcriptome in human brain. *Nature Neurosci.* 11, 1271–1282 (2008).
10. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4, 17 (2005).
11. Acampora, D. *et al.* Craniofacial, vestibular and bone defects in mice lacking the Distal-less-related gene Dlx5. *Development* 126, 3795–3809 (1999).
12. Yoshihara, M., Adolfsen, B., Galle, K. T. & Littleton, J. T. Retrograde signaling by Syt 4 induces presynaptic release and synapse-specific growth. *Science* 310, 858–863 (2005).
13. Brusse, E. *et al.* Spinocerebellar ataxia associated with a mutation in the fibroblast growth factor 14 gene (SCA27): A new phenotype. *Mov. Disord.* 21, 396–401 (2006).
14. Holmes, S. E. *et al.* Expansion of a novel CAG trinucleotide repeat in the 5′ region of PPP2R2B is associated with SCA12. *Nature Genet.* 23, 391–392 (1999).
15. Belton, E., Salmond, C. H., Watkins, K. E., Vargha-Khadem, F. & Gadian, D. G. Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Hum. Brain Mapp.* 18, 194–200 (2003).
16. Shu, W. *et al.* Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene. *Proc. Natl Acad. Sci. USA* 102, 9643–9648 (2005).
17. Van Camp, G. *et al.* A new autosomal recessive form of Stickler syndrome is caused by a mutation in the COL9A1 gene. *Am. J. Hum. Genet.* 79, 449–457 (2006).
18. Uhlenberg, B. *et al.* Mutations in the gene encoding gap junction protein α12 (connexin 46.6) cause Pelizaeus-Merzbacher-like disease. *Am. J. Hum. Genet.* 75, 251–260 (2004).
19. Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* 62, 494–509 (2009).
20. Prabhakar, S., Noonan, J. P., Paabo, S. & Rubin, E. M. Accelerated evolution of conserved noncoding sequences in humans. *Science* 314, 786 (2006).
21. Dorus, S. *et al.* Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. *Cell* 119, 1027–1040 (2004).
22. Enard, W. *et al.* A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell* 137, 961–971 (2009).
23. Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* 392, 917–920 (1998).
24. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* 188, 107–116 (1975).
25. Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* 296, 340–343 (2002).
26. Caceres, M. *et al.* Elevated gene expression levels distinguish human from non-human primate brains. *Proc. Natl Acad. Sci. USA* 100, 13030–13035 (2003).
27. Coppola, G. *et al.* Gene expression study on peripheral blood identifies progranulin mutations. *Ann. Neurol.* 64, 92–96 (2008).

# METHODS

**Antibodies.** The following antibodies were either used for immunoblotting (IB) or immunofluorescence (IF): anti-Flag (mouse monoclonal, Sigma; 1:10,000 (IB), 1:10,000 (IF)), anti-GAPDH (mouse monoclonal, Chemicon; 1:2500 (IB)), anti-β-tubulin (rabbit polyclonal, Abcam; 1:1000 (IB)), anti-FOXP1 (ref. 6; 1:5000 (IB), 1:1000 (IF)), anti-CACNB2 (mouse monoclonal, Abcam; 1:100 (IB)), anti-ENPP2 (rabbit polyclonal, Cayman Chemical; 1:400 (IB)), goat anti-rabbit horseradish peroxidase (Cell Signaling, 1:2,500), goat anti-mouse horseradish peroxidase (Chemicon, 1:5,000), goat anti-mouse Alexa Fluor 488 (Invitrogen, 1:1,500), goat anti-rabbit Alexa Fluor 594 (Invitrogen, 1:1,500).

**Cell culture and stable line generation.** Stable SH-SY5Y cell lines were generated by transfecting cells with pCMV-Tag4a expression constructs using FuGENE (Roche Applied Science) according to the manufacturer's instructions. Populations of stable cells were selected using $1\,mg\,ml^{-1}$ geneticin (Invitrogen). Multiple independent lines were generated from independent transfections. Stable human fetal neuronal progenitor cell lines were generated by transducing cells with lentiviruses as previously described[28]. FOXP2-producing lentiviral vectors were generated by replacing the eGFP in pLUGIP (ATCC) with FOXP2.

**Immunoprecipitation.** Nuclear extract was incubated with either 1 µg of Flag antibody (Sigma) or a polyclonal FOXP1 antibody[6].

**Cell proliferation assay.** Equal numbers of cells ($2.0 \times 10^4$) were plated on time zero and counted every subsequent day after trypsinization using a haemacytometer.

**Dual luciferase assays.** 293T cells (ATCC) were transfected with 50 ng of reporter construct expressing *Photinus pyralis* (firefly) luciferase, 1 ng of *Renilla* luciferase plasmid (pRL-EF), and 50 ng of pCMV-Tag4a FOXP2 expression plasmid using FuGENE (Roche Applied Science) according to the manufacturer's instructions. Forty-eight hours later, cells were lysed and analysed using the dual luciferase reporter assay system (Promega) according to the manufacturer's instructions. Co-transfection of Renilla was used for transfection normalization, and values were additionally normalized to cells transfected with a promoter-less luciferase construct. Promoter information is in Supplementary Table 6. The canonical FOXP2 binding site driving luciferase was generated by cloning AATTTG in triplicate into pGL4 (Promega).

**Gene ontology analysis.** GO analysis was performed as described[6] using DAVID (http://david.abcc.ncifcrf.gov). The differentially expressed genes were compared to all of the genes on the microarrays and a *P* value computed using a Fisher's exact test.

**Immunoblotting.** Whole-cell protein lysates were generated and immunoblotted as described[28].

**Immunofluorescence.** Cells were grown on glass coverslips, fixed in 2% paraformaldehyde, and permeabilized in 0.2% Triton X-100. TBST containing 10% milk and 10% normal goat serum was used as blocking solution at room temperature for 1 h. Antibodies were diluted in TBS with 0.25% BSA, 0.25% normal goat serum and 0.1% Triton X-100 and applied to cells overnight at 4 °C. Secondary antibodies were diluted in blocking solution and added at room temperature for 1 h. Coverslips were mounted to glass slides and images taken using a Zeiss Axio Imager D1.

**Mass spectrometry.** FOXP2 immunoprecipitates were precipitated by the addition of trichloroacetic acid and proteolysed by the sequential addition of Lys-C and trypsin proteases[29]. Digested peptide samples were then analysed by mass spectrometry as described[29]. Proteins were considered to be present in a sample if at least two peptides per protein were identified using a false positive rate of less than 5% per peptide as determined by a decoy database strategy[30].

**Microarrays.** For the SH-SY5Y data, we analysed four biological replicates of each genotype from three independently generated cell lines for a total of 12 microarrays per genotype. Each of these cell lines was created from populations of cells rather than single clones, and as such, the expression data represent changes from hundreds of independent integrations throughout the cells' genomes. Furthermore, as the endogenous FOXP2 expression is very low in SH-SY5Y cells, the potential confound of heterodimerization with endogenous human FOXP2 is mitigated in these cells. For the tissue data, we analysed three to six independent samples for each brain region in each species. Detailed code for the microarray analysis is available[31].

**Permutation testing.** For FOXP2 correlations, we computed the average correlation for each gene on the microarray to either the level of the human or the chimpanzee FOXP2. We then derived the absolute difference in correlation for each gene between the human and chimpanzee FOXP2 arrays. The average of these differences was not statistically different from performing the same test while randomizing the correlation values for all of the genes on the arrays, or using the values from only the differentially expressed genes. For promoter binding site calculations, we calculated the number of promoters from differentially expressed genes with a given motif and compared them to the average number from a random selection of the same number of promoters from the genome. We assumed a normal distribution and a *Z*-score less than 0.05 was called significant. Similar analysis was done for comparing genes with a haCNS and expression in human fetal brain. For microarray overlap comparisons, we included the number of differentially expressed genes as well as the total number of probe sets on the microarrays for each comparison. We used a hypergeometric distribution test with 10,000 permutations to calculate the mean and standard deviation of the overlap. We assumed a normal distribution, and a *Z*-score less than 0.05 was called significant.

**Real-time PCR.** RNA extraction and RT–PCR was performed as described[6]. Primer sequences are in Supplementary Table 10.

**Site-directed mutagenesis.** Mutagenesis of pCMV-Tag4a/FOXP2[6] was carried out using the GeneTailor Site-Directed Mutagenesis System (Invitrogen) according to the manufacturer's instructions using the following primers: site 1 (asparagine to threonine), F-5′-CCTCCTCGACTACCTCCTCCACAACTTCC AAAGC-3′, R-5′-GGAGGAGGTAGTCGAGGAGGAATTGTTAGTA-3′; site 2 (serine to asparagine), F-5′-ATGGACAGTCTTCAGTTCTAAACGCAAGACG AGA-3′, R-5′-TAGAACTGAAGACTGTCCATTCACTATGGAA-3′. Mutagenesis was confirmed by both sequencing and mass spectrometry.

**Weighted gene coexpression network analysis (WGCNA).** WGCNA was performed as previously described[9,10]. Briefly, genes were chosen for inclusion into the network on the basis of their consistent presence on the array and high coefficient of variation, and they were clustered based on their topological overlap. For each module, singular value decomposition ($X = UDV'$) was performed, and the expression was re-calculated without the first principal component because it corresponded to cell line differences. The modules reported in this study were created using expression data with the first principal component removed, as it represented an experimental batch effect.

28. Konopka, G., Tekiela, J., Iverson, M., Wells, C. & Duncan, S. A. Junctional adhesion molecule-A is critical for the formation of pseudocanaliculi and modulates E-cadherin expression in hepatic cells. *J. Biol. Chem.* **282,** 28137–28148 (2007).

29. Wohlschlegel, J. A. Identification of SUMO-conjugated proteins and their SUMO attachment sites using proteomic mass spectrometry. *Methods Mol. Biol.* **497,** 33–49 (2009).

30. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* **4,** 207–214 (2007).

31. Coppola, G., Winden, K., Konopka, G., Gao, F. & Geschwind, D. H. Expression and network analysis of Illumina microarray data. *Nature Protocols* doi:10.1038/nprot.2009.215 (2009).

# LETTERS

# Bidirectional plasticity in fast-spiking GABA circuits by visual experience

Yoko Yazaki-Sugiyama[1,2], Siu Kang[3], Hideyuki Câteau[3], Tomoki Fukai[3] & Takao K. Hensch[1,2]†

Experience-dependent plasticity in the brain requires balanced excitation–inhibition[1]. How individual circuit elements contribute to plasticity outcome in complex neocortical networks remains unknown. Here we report an intracellular analysis of ocular dominance plasticity—the loss of acuity and cortical responsiveness for an eye deprived of vision in early life[2,3]. Unlike the typical progressive loss of pyramidal-cell bias, direct recording from fast-spiking cells *in vivo* reveals a counterintuitive initial shift towards the occluded eye followed by a late preference for the open eye, consistent with a spike-timing-dependent plasticity rule for these inhibitory neurons. Intracellular pharmacology confirms a dynamic switch of GABA (γ-aminobutyric acid) impact to pyramidal cells following deprivation in juvenile mice only. Together these results suggest that the bidirectional recruitment of an initially binocular GABA circuit may contribute to experience-dependent plasticity in the developing visual cortex.

For nearly 50 years, the response to visual stimulation of one or the other eye (ocular dominance) has been the leading model of cortical processing and plasticity[2], and is known to require mature inhibitory circuit function[4]. Imaging efforts to monitor plasticity within the heterogeneous GABAergic population have led to inconsistent results, including reports of little or delayed change in responsiveness[5,6]. We have applied intracellular recording techniques to the binocular zone of mouse primary visual cortex *in vivo* (Fig. 1a), allowing both morphological and physiological identification, as well as slow pharmacological access from within the same cell. Increasing evidence notably implicates a single cell type—the parvalbumin-positive large basket cell—to be particularly important for visual cortical plasticity[1,7–10].

These local interneurons are readily identified by their characteristic thin spike shape (mean half-width, $1.03 \pm 0.07$ ms) and prominent after-hyperpolarizations (Fig. 1a, inset trace), as well as post hoc morphological visualization (Fig. 1b). Unexpectedly, we found that such fast-spiking interneurons exhibit a dynamic profile of ocular dominance plasticity distinct from the classic shift observed in pyramidal cells[2]. The spike response of the majority of pyramidal cells in mice is normally biased in favour of the contralateral eye and shifts progressively towards
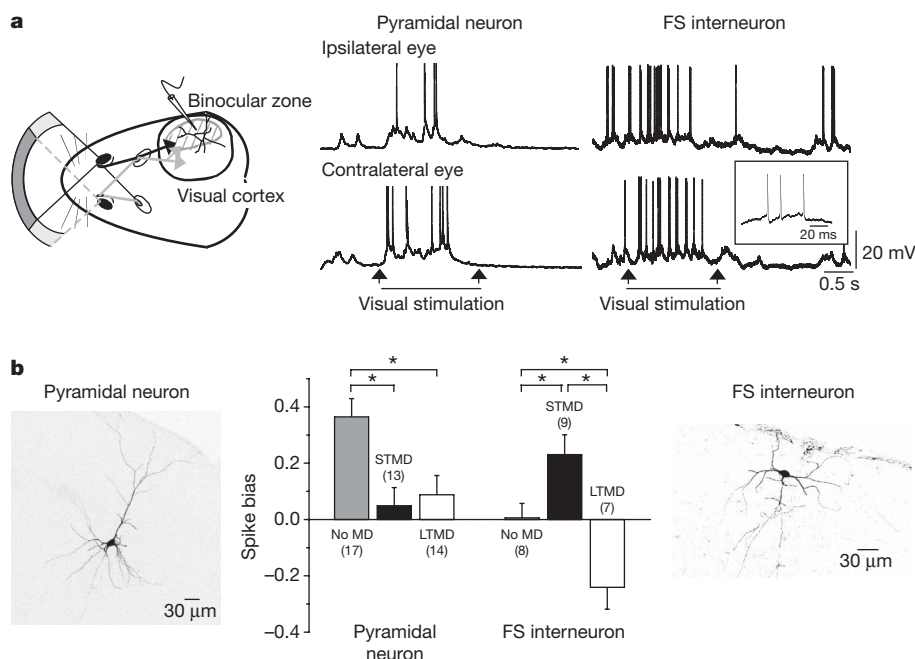


**Figure 1 | Distinct plasticity profiles in pyramidal and fast-spiking visual cortical neurons. a,** Schematic of intracellular recording from mouse visual cortex (left) and sample pyramidal-cell and fast-spiking (FS)-cell response to visual stimulation of both eyes (middle and right). Inset, magnified fast spike shape. **b,** Micrographs of a dye-filled pyramidal cell (left) and a fast-spiking interneuron (right). The central plot shows spike bias (mean ± s.e.m.) as a function of cell type in mice respectively undergoing no monocular deprivation (no MD), STMD and LTMD at P24–P29. The number of cells per group is indicated. We note the counterintuitive shift in fast-spiking cells initially towards the deprived eye (no-MD bias less than STMD bias) then towards the open eye (no-MD bias greater than LTMD bias). *$P < 0.05$, unpaired *t*-test.

[1]CREST, JST, Toyonaka, Osaka 560-0082, Japan. [2]Laboratory for Neuronal Circuit Development, [3]Laboratory for Neural Circuit Theory, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. †Present address: Center for Brain Science, Harvard University, 52 Oxford Street, Cambridge, Massachusetts 02138, USA.

the open eye following both short-term (three-day) and long-term (two-week) monocular deprivation (STMD and LTMD, respectively) (Fig. 1b, left)[4]. In striking contrast, fast-spiking interneurons normally displayed no ocular dominance bias, responding equally well to both eyes. However, upon STMD they first exhibited a counterintuitive shift in favour of the contralateral, deprived, eye before acquiring an open-eye bias after LTMD (Fig. 1b, right).

Robust bidirectional ocular dominance plasticity in fast-spiking interneurons challenges the view that GABA cells may lack plasticity mechanisms[11]. Recently, a spike-timing-dependent plasticity (STDP) rule was proposed for intracortical connections to a fast-spiking interneuron subtype in slices of somatosensory cortex[12]. We have tested whether such STDP is sufficient to explain bidirectional ocular dominance plasticity in fast-spiking interneurons of visual cortex *in vivo*. We constructed a model in which a GABAergic neuron receives excitatory inputs from both the ipsilateral and the contralateral eye, as well as inhibitory inputs to maintain firing rates within a physiological range (Fig. 2a, top).

A conductance bias shift initially favouring the deprived eye and later the open eye was recapitulated by applying the bilaterally symmetric STDP rule[12] (Fig. 2a, bottom) in conjunction with an additional slow elimination of inactive synapses (Fig. 2b). Because STDP weakens active connections more than inactive ones, an extra mechanism to remove inactive synapses is needed to achieve the eventual reversal of ocular dominance bias after prolonged monocular deprivation (Supplementary Fig. 1). Consistent with such pruning, our recordings *in vivo* revealed a striking reduction of spontaneous activity within fast-spiking cells upon monocular deprivation (Fig. 2c). This was not seen in pyramidal cells (no monocular deprivation, $2.0 \pm 0.5$ Hz; STMD, $2.1 \pm 0.4$ Hz; LTMD, $1.3 \pm 0.4$ Hz), which do not undergo bidirectional ocular dominance changes (Fig. 2c).

We next examined how this dynamic change in the function of inhibitory circuits is reflected in their pyramidal-cell targets. Optimal activation of perisomatic GABA$_A$ receptors containing the α1-subunit are essential for ocular dominance plasticity to occur[7,8]. We therefore probed the influence of GABA function within single

pyramidal cells by taking advantage of our direct intracellular access. Sharp microelectrodes can slowly deliver dyes or reagents from the cell body outward into the dendrites (Fig. 1b, micrographs), allowing direct physiological comparison before and after drug application *in vivo*. We first determined whether the extended recording duration required for intracellular diffusion would alter responsiveness or cause the drug to leak out of the cell. Prolonged intracellular recordings ($>30$ min) either with no drug or with a small, competitive GABA$_A$-receptor antagonist (1 mM gabazine) expected to act only on the extracellular amino terminus failed to produce a change in pyramidal-cell spike bias regardless of deprivation status (Fig. 3a).

In contrast, intracellular recording *in vivo* with picrotoxin (1 mM PTX) significantly altered spike bias 30 min after impalement (Fig. 3b). This classic non-competitive antagonist has successfully been used to block the pores of open GABA$_A$-receptor channels from the intracellular side of cortical neurons[13–16]. In slices of mouse visual cortex, we confirmed that PTX (1 mM) internal solutions eliminate GABA$_A$-mediated miniature inhibitory postsynaptic currents (A. Scheuber and T.K.H., unpublished observations; Supplementary Fig. 2), which largely arise from perisomatic inputs[17]. Consistent with this, we noticed a subtle (statistically insignificant) reduction of resting membrane potential and firing threshold under our conditions *in vivo* ($P > 0.05$, paired *t*-tests; Supplementary Fig. 3a). Notably, the biased spike response in normal juvenile mice was robustly lost by applying PTX within single pyramidal neurons ($P < 0.05$, paired *t*-tests; Fig. 4a, 'No MD').

Remarkably, after STMD during the peak of the critical period, P24–P29 (P, postnatal day)[4,18], the same intracellular PTX infusion had a very different effect (Supplementary Fig. 4). The initial ocular dominance spike bias was consistently inverted within each cell
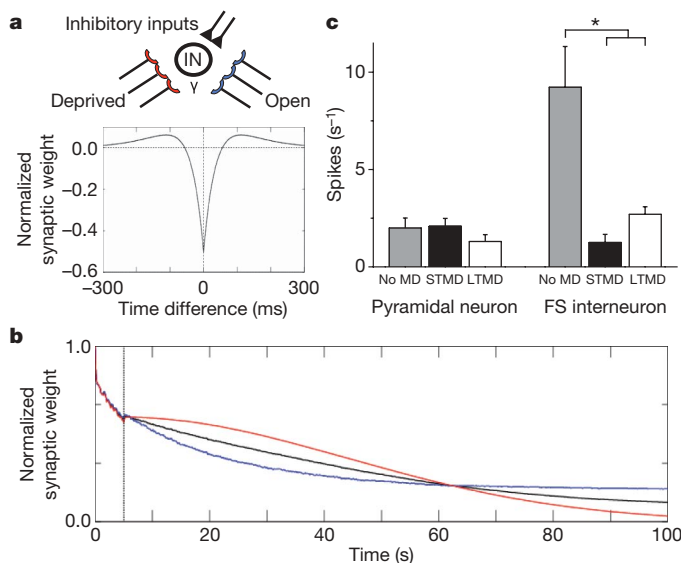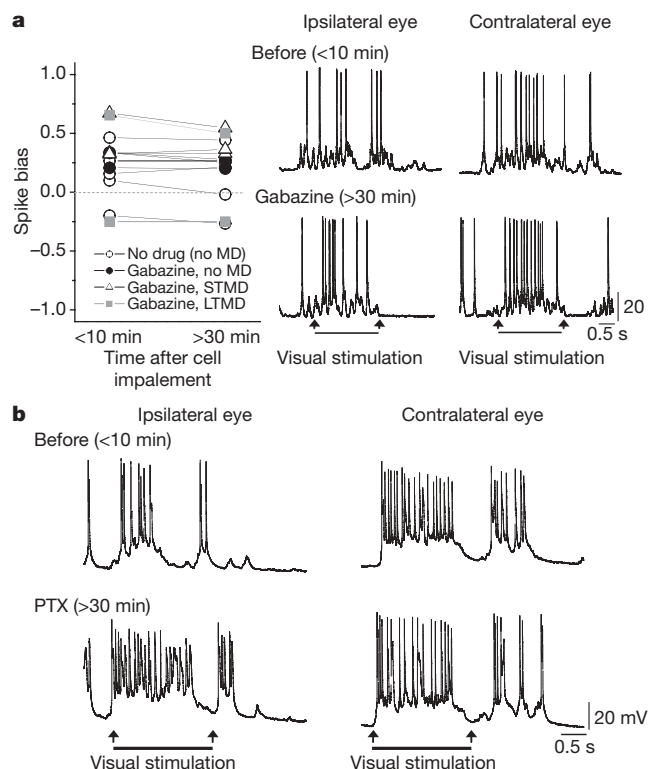
**Figure 2 | Bidirectional plasticity in fast-spiking cells *in vivo* follows an STDP rule. a**, Model fast-spiking cell (top) receiving inhibitory input to limit firing rate (with no impact on plasticity outcome) and excitatory visual inputs observing an STDP rule (bottom)[12]. **b**, Simulation time course of average synaptic weight from deprived eye (red), open eye (blue) and both eyes (black). Deprived-eye firing rate was decreased from 20 to 0.2 Hz at $t = 5$ s. Conductance biases invert at $t = 63$ s (Supplementary Fig. 1a). **c**, Spontaneous firing rate *in vivo* decreases significantly only in fast-spiking cells upon deprivation (no MD, $9.2 \pm 2.1$ Hz; STMD, $1.3 \pm 0.2$ Hz; LTMD, $2.7 \pm 0.4$ Hz; *P < 0.05, analysis of variance).

**Figure 3 | Intracellular GABA-receptor blockade in pyramidal neurons of visual cortex. a**, Spike bias remains unchanged (for $>30$ min) before and after intracellular gabazine application (no MD, filled circles; STMD, triangles; LTMD, squares) or no drug application (open circles) to single pyramidal neurons (left). Sample traces from the same cell before and after intracellular gabazine application are shown on the right. **b**, Sample traces before and after intracellular picrotoxin (1 mM PTX) application ($>30$ min).
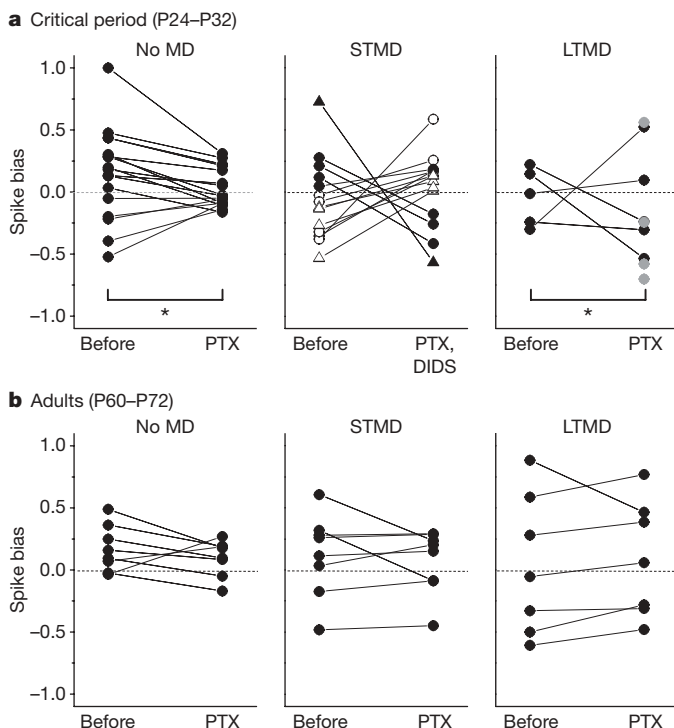
**Figure 4 | Dynamic GABA impact in single pyramidal cells with age and experience. a**, Critical-period (P24–P29) spike bias in single pyramidal neurons before and after intracellular GABA-receptor blockade is reduced by PTX (no MD). After STMD, PTX (circles) or 4,4′-diisothiocyanato-stilbene-2,2′-disulphonic acid (DIDS, triangles) invert ipsilateral-biased cells (open symbols) to prefer contralateral input and contralateral-biased cells (closed symbols) to prefer ipsilateral input ($P < 0.001$, binomial test). Some LTMD cells acquired a strongly biased spike response only after PTX (grey symbols). **b**, Adult (P60–P72) spike bias remains largely unaltered before and after intracellular PTX application. *$P < 0.01$, paired $t$-tests.

(Fig. 4a, STMD circles). This unexpected flipping of ocular dominance bias after STMD was confirmed using another intracellular GABA$_A$-receptor blocker, DIDS[16,19] (Fig. 4a, STMD triangles). Such an inversion effect of inhibition would have been overlooked by extracellular measures of population activity, as there was no absolute change in mean ocular dominance bias before and after infusion of PTX ($P > 0.6$, paired $t$-tests). The ocular dominance inversion effect of PTX persisted even as ocular dominance plasticity became saturated by deprivation for two weeks or more[6,18] (Fig. 4a, LTMD). At that time, some cells that received only a subthreshold depolarization before drug exposure acquired a strong biased spike response after PTX infusion (LTMD grey circles). Systemic application of GABA$_A$ antagonists (for example bicuculline) has previously been suggested to unmask latent deprived-eye input following LTMD in kittens[20].

Using network simulations, we further examined whether the dynamic inhibitory changes within single pyramidal neurons could be explained by our initially observed interneuron plasticity (Fig. 1b). We constructed a minimal model of the cortical circuit and tested various combinations of either plastic or fixed synapses at three different sites (Supplementary Fig. 5a). Because plastic GABAergic synapses onto pyramidal neurons in monocular visual cortex have been reported recently[21], we further tested whether monocular plastic GABAergic synapses onto pyramidal cells could account for the experimental observation. Of these four models, only those with plastic excitatory synapses onto both pyramidal cells and interneurons (Supplementary Fig. 5a, leftmost diagram) or those with plastic monocular GABAergic synapses (Supplementary Fig. 5a rightmost diagram) could successfully reproduce the experimentally observed dynamics (Supplementary Fig. 5b, c). In the first model, we also

observed bidirectional conductance changes of excitatory input onto interneurons, first biased towards the deprived eye and then towards the open eye (Supplementary Fig. 5d).

Notably, in the binocular zone where ocular dominance competition actually occurs *in vivo*, we have yet to observe purely monocular fast-spiking cells (Fig. 1b), indicating that a bidirectional recruitment model (Supplementary Fig. 5a, leftmost diagram) is most likely. To determine whether such dynamic reconfiguration of inhibitory function reflects a true critical-period plasticity, we also conducted the same PTX experiment in adult mice (>P60), which do not lose visual input or acuity upon monocular deprivation[3,18]. Intracellular PTX had little or no effect on initial spike bias and no ocular dominance inversion was observed after monocular deprivation (Fig. 4b). This indicates that the inhibitory circuit is particularly plastic only during a critical period and further rules out a general, non-specific action of PTX on cellular physiology. With age, fast-spiking interneurons are increasingly enwrapped by tight perineuronal nets, which may structurally limit plasticity of their afferent inputs[22].

Thus, direct intracellular recording corrects the erroneous view that parvalbumin neurons are particularly refractory to deprivation, which is based largely on measures of non-spiking activity, such as c-Fos labelling[6]. Recent two-photon imaging of genetically labelled GABA cells supports our result insofar as all GABA cells may shift towards the open eye after one week of monocular deprivation[5]. However, an initial counterintuitive shift was not observed previously. Imaging *in vivo* is necessarily biased to superficial layers that contain fewer fast-spiking cells, and GAD67–GFP knock-in mice also suffer a reduction of GABA content known to impair early stages of experience-dependent plasticity[4].

As specific inhibitory cell types are involved in ocular dominance plasticity[1], detailed cell identification is essential. Indeed, fast-spiking interneurons receive the most potent and direct thalamocortical input[23,24]. Synchronized, fast and broadly tuned feed-forward inhibition is well suited to the suppression of spike generation in pyramidal cells following all but the most optimal excitatory input—narrowing a brief 'window of excitability' for plasticity. In the visual cortex, circuit rewiring then follows two temporally and mechanistically distinct phases: an early loss of deprived-eye response followed by a delayed homeostatic increase of open-eye input[25]. The early phase is suggested to involve an input-specific synaptic plasticity[26], but it remains unclear to what extent long-term depression of excitatory postsynaptic currents occurs in pyramidal neurons following monocular deprivation[21].

Our findings offer a novel substrate for the early phase of the ocular dominance plasticity process. Enhanced activation of inhibitory fast-spiking cells by deprived-eye afferents may initially suppress their own impact on pyramidal cells. The strengthening of inhibitory drive in a direction initially opposed to excitatory network adjustment is reminiscent of early stages of barn owl adaptive map refinement[27] and may reflect a cortical manifestation *in vivo* of anti-Hebbian plasticity rules recently described for hippocampal interneurons *in vitro*[28]. Importantly, fast-spiking interneurons are selectively modulated by visual attention[29], an essential condition for correcting amblyopia[30]. Understanding mechanisms of plasticity in fast-spiking cells promises deeper insight into disorders of critical-period brain development.

## METHODS SUMMARY

***In vivo* intracellular recording.** We performed *in vivo* intracellular recording from layer 2/3 of the binocular visual cortex in mice (P24–P32; strain C57Bl/6) with or without monocular deprivation (contralateral to the deprived eye). One eyelid was sutured under anaesthesia either two to three days (STMD) or more than two weeks (LTMD) before recording. In some cases, adult mice (P60–P72) were used. Mice were anaesthetized and fixed to a stereotaxic apparatus[4,18] and a sharp glass electrode (90–160 MΩ; 2 M potassium acetate, 5% neurobiotin) was inserted into the brain. To block all GABA-receptor currents, some electrodes were filled with picrotoxin (1 mM PTX in 2 M caesium acetate). Additional prolonged recordings were performed either with gabazine (1 mM in 2 M potassium acetate) or with no drug-filled electrodes to control for PTX specificity.

Intracellular membrane potentials (mean ± s.e.m.: pyramidal cell, $-79 \pm 3$ mV; fast-spiking cell, $-69 \pm 2$ mV) were recorded in current-clamp mode (Axoclamp-2B, Axon Instruments). Neurons were filled with neurobiotin for post hoc anatomical identification. Collected data were analysed (using MATLAB, version 7.0.4) for spike output. Response strength (firing rate during visual stimulus minus baseline firing rate) was calculated for each eye and an ocular dominance bias index, given by (contralateral-eye response − ipsilateral-eye response)/(contralateral-eye response + ipsilateral-eye response), was calculated for the average of seven to twenty stimuli applied to each eye. To assess any drug effect, we compared ocular dominance bias indices obtained 3–10 min after cell penetration with those obtained 30–45 min later. Five to fifteen responses were averaged for each time point.

**Computational modelling.** For STDP modelling, we adapted a simple fast-spiking-interneuron model that receives 100 excitatory Poissonian inputs of 20 Hz from both the contralateral and the ipsilateral eye. Forty inhibitory inputs of 10 Hz were also introduced to regulate the firing rate. The STDP rule was applied at excitatory synapses.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Hensch, T. K. Critical period plasticity in local cortical circuits. *Nature Rev. Neurosci.* **6**, 877–888 (2005).
2. Wiesel, T. N. & Hubel, D. H. Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J. Neurophysiol.* **26**, 1003–1017 (1963).
3. Prusky, G. T. & Douglas, R. M. Developmental plasticity of mouse visual acuity. *Eur. J. Neurosci.* **17**, 167–173 (2003).
4. Hensch, T. K. *et al.* Local GABA circuit control of experience-dependent plasticity in developing visual cortex. *Science* **282**, 1504–1508 (1998).
5. Gandhi, S. P., Yanagawa, Y. & Stryker, M. P. Delayed plasticity of inhibitory neurons in developing visual cortex. *Proc. Natl Acad. Sci. USA* **105**, 16797–16802 (2008).
6. Mainardi, M., Landi, S., Berardi, N., Maffei, L. & Pizzorusso, T. Reduced responsiveness to long-term monocular deprivation of parvalbumin neurons assessed by c-Fos staining in rat visual cortex. *PLoS ONE* **4**, e4342 (2009).
7. Fagiolini, M. *et al.* Specific GABA-A circuits for visual cortical plasticity. *Science* **303**, 1681–1683 (2004).
8. Katagiri, H., Fagiolini, M. & Hensch, T. K. Optimization of somatic inhibition at critical period onset. *Neuron* **53**, 805–812 (2007).
9. Di Cristo, G. *et al.* Activity-dependent PSA expression regulates inhibitory maturation and onset of critical period plasticity. *Nature Neurosci.* **10**, 1569–1577 (2007).
10. Sugiyama, S. *et al.* Experience-dependent transfer of Otx2 homeoprotein into the visual cortex activates postnatal plasticity. *Cell* **134**, 508–520 (2008).
11. McBain, C. J., Freund, T. F. & Mody, I. Glutamatergic synapses onto hippocampal interneurons: precision timing without lasting plasticity. *Trends Neurosci.* **22**, 228–235 (1999).
12. Lu, J. T., Li, C. Y., Zhao, J. P., Poo, M. M. & Zhang, X. H. Spike-timing-dependent plasticity of neocortical excitatory synapses on inhibitory interneurons depends on target cell type. *J. Neurosci.* **27**, 9711–9720 (2007).
13. Inomata, N., Tokutomi, N., Oyama, Y. & Akaike, N. Intracellular picrotoxin blocks pentobarbital-gated Cl- conductance. *Neurosci. Res.* **6**, 72–75 (1988).
14. Metherate, R. & Ashe, J. H. Ionic flux contributions to neocortical slow waves and nucleus basalis-mediated activation: whole-cell recordings *in vivo. J. Neurosci.* **13**, 5312–5323 (1993).
15. Mancilla, J. G. & Ulinski, P. S. Role of GABA(A)-mediated inhibition in controlling the responses of regular spiking cells in turtle visual cortex. *Vis. Neurosci.* **18**, 9–24 (2001).
16. Nelson, S., Toth, L., Sheth, B. & Sur, M. Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science* **265**, 774–777 (1994).
17. Soltesz, I., Smetters, D. K. & Mody, I. Tonic inhibition originates from synapses close to the soma. *Neuron* **14**, 1273–1283 (1995).
18. Gordon, J. A. & Stryker, M. P. Experience-dependent plasticity of binocular responses in the primary visual cortex of the mouse. *J. Neurosci.* **16**, 3274–3286 (1996).
19. Dudek, S. M. & Friedlander, M. J. Intracellular blockade of inhibitory synaptic responses in visual cortical layer IV neurons. *J. Neurophysiol.* **75**, 2167–2173 (1996).
20. Duffy, F. H., Burchfiel, J. L. & Conway, J. L. Bicuculline reversal of deprivation amblyopia in the cat. *Nature* **260**, 256–257 (1976).
21. Maffei, A., Nataraj, K., Nelson, S. B. & Turrigiano, G. G. Potentiation of cortical inhibition by visual deprivation. *Nature* **443**, 81–84 (2006).
22. Pizzorusso, T. *et al.* Reactivation of ocular dominance plasticity in the adult visual cortex. *Science* **298**, 1248–1251 (2002).
23. Cruikshank, S. J., Lewis, T. J. & Connors, B. W. Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex. *Nature Neurosci.* **10**, 462–468 (2007).
24. Erisir, A. & Dreusicke, M. Quantitative morphology and postsynaptic targets of thalamocortical axons in critical period and adult ferret visual cortex. *J. Comp. Neurol.* **485**, 11–31 (2005).
25. Kaneko, M., Stellwagen, D., Malenka, R. C. & Stryker, M. P. Tumor necrosis factor-alpha mediates one component of competitive, experience-dependent plasticity in developing visual cortex. *Neuron* **58**, 673–680 (2008).
26. Frenkel, M. Y. & Bear, M. F. How monocular deprivation shifts ocular dominance in visual cortex of young mice. *Neuron* **44**, 917–923 (2004).
27. Zheng, W. & Knudsen, E. I. Gabaergic inhibition antagonizes adaptive adjustment of the owl's auditory space map during the initial phase of plasticity. *J. Neurosci.* **21**, 4356–4365 (2001).
28. Lamsa, K. P., Heeroma, J. H., Somogyi, P., Rusakov, D. A. & Kullmann, D. M. Anti-Hebbian long-term potentiation in the hippocampal feedback inhibitory circuit. *Science* **315**, 1262–1266 (2007).
29. Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* **55**, 131–141 (2007).
30. Levi, D. M. & Li, R. W. Improving the performance of the amblyopic visual system. *Phil. Trans. R. Soc. Lond. B* **364**, 399–407 (2009).

**Author Contributions** T.K.H. and Y.Y.-S. designed the experiments; S.K., H.C. and T.F. formulated the computational models; Y.Y.-S. performed the intracellular recording *in vivo*; S.K. ran the simulations; and T.K.H., Y.Y.-S., S.K. and T.F. wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.K.H. (hensch@mcb.harvard.edu).

## METHODS

**Animals and monocular deprivation.** We performed *in vivo* intracellular recording in visual cortex of mice (P24–P32, strain C57Bl/6). For STMD, one eyelid was sutured under halothane anaesthesia two to three days before recording. For LTMD, the eyelid was sutured at P18 for a duration of more than two weeks[1,4]. In some cases, adult wild-type mice (P60–P72, strain C57Bl/6) were used. For STMD in adult mice, one eyelid was sutured two to three days before recording, and for LTMD, sutures were placed at P49–P56 for more than two weeks.

**Intracellular recording and data analysis.** Mice were anaesthetized with pentobarbital and fixed to the stereotaxic apparatus after tracheotomy, as described previously[3,5]. Supplemental halothane inhalation during recording was added as needed. A small opening was made above the monocular zone in the skull (secured by dental cement), and a sharp glass electrode was inserted into the brain at an oblique (30–35°) angle to stably record from layer-2/3 neurons of the binocular visual cortex. Intracellular membrane potentials were recorded in current-clamp mode (Axoclamp-2B, Axon Instruments). All recordings were obtained in bridge mode, which was balanced by offsetting the electrode series resistance to hyperpolarizing current pulse injections (350 pA, 50 ms) before the cell was impaled.

The bridge was checked periodically and rebalanced as needed during recording. Cells with a resting membrane potential more negative than −60 mV and robust spontaneous synaptic activity were probed with visual stimuli. In some cases, a small negative tonic current was applied through the electrode to hyperpolarize the membrane potentials to accentuate the evoked synaptic or action-potential activity. Response comparisons with stimulation of either eye were made at the same membrane potentials. Intracellular potentials were recorded, digitized (20 kHz), stored on a personal computer and analysed (using DataWave Technologies SCIWORKS software, version 3) offline. Neurons were filled with neurobiotin after recording and identified by post hoc histology[31].

Five to twenty sweeps of visual stimulation (90° moving light bar; 100% contrast; sweep speed, 4° s$^{-1}$) were driven by a computer and projected onto a tangent screen in front of the animal (25–30 cm; 30° angle). The eyes were alternately stimulated at 5-s intervals using remote-controlled motorized eye shutters (serial no. 981202, Optical Imaging). Visual stimulation was presented to each eye for 6 s and recorded for 10 s, including a non-stimulus baseline period before and after each visual stimulus. A significant response above baseline activity was determined by paired *t*-tests ($P < 0.05$). The response strength in spike response was calculated as RS = SFR − BFR, where SFR is the firing rate during stimulus presentation and BFR is the firing rate during a post-stimulus baseline period of equal duration. Average response strength was calculated for 7–20 stimulus sweeps to each eye, and the ocular dominance bias index was calculated[32].

To determine the drug effect statistically, we used the paired *t*-tests and the binomial test (Fig. 4). To test for a decrease or increase of bias, we applied paired *t*-tests ($P < 0.05$) between the absolute value of bias index before and after drug application within the same cells. To see an inversion of the bias (Fig. 4a, STMD), we calculated the changing probability of positive and negative signs in the bias index using the binomial test ($P < 0.05$).

**Local circuit model: STDP model.** The fast-spiking interneuron modelled in ref. 33 was used in the STDP model (Fig. 2a). The neuron received 100 excitatory Poissonian inputs of 20 Hz from both the contralateral and the ipsilateral eye. Forty inhibitory inputs of 10 Hz were also introduced to regulate the firing rate. Synaptic conductances were chosen to be $g_{AMPA,max} = 0.03$ mS cm$^{-2}$ and $g_{GABA} = 0.1$ mS cm$^{-2}$. The STDP rule at the excitatory synapses was taken from ref. 12. All excitatory synaptic conductances were modified as follows: $g \rightarrow g + G(t)g_{max}$, where $t = t_{post} - t_{pre}$, $g_{max}$ is the maximum synaptic conductance and $G(t) = A\exp(-|t|/\tau_A) + B\exp(-|t|/\tau_A)$. The parameters were chosen to be $A = -0.01$, $B = 0.00495$, $\tau_A = 40$ ms and $\tau_B = 80$ ms. Positive or negative $G(t)$ implies long-term potentiation (LTP) or long-term depression (LTD), respectively; our parameter choices result in a LTP/LTD area ratio of 0.99. At the onset of monocular deprivation, the presynaptic firing rate was reduced to 0.2 Hz at the deprived-eye synapses and the elimination of the inactive synapses was initiated. The details of the latter process remain unknown, but it should be slower than the modifications by STDP. Here, the process was modelled by a multiplicative operation, $g(t) \rightarrow g(t)s(t)$, at each time step of the simulations, where the multiplicative factor obeys $ds/dt = -s/\tau_s$ with $s(0) = 1$ and $\tau_s = 6$ months.

**Local circuit model: network organization.** We also calculated spike bias (mean ± s.e.m.) for our local circuit model. Our network model consisted of a four-compartment pyramidal neuron and an inhibitory interneuron. We have

also tested models using ten inhibitory neurons, but the results did not essentially depend upon the number of interneurons. The four-compartment pyramidal neuron (comprising soma and basal, proximal and distal dendrites) received excitatory afferent inputs from ipsilateral and contralateral eyes and inhibitory input from a single-compartment interneuron. The afferent inputs terminated on the distal dendrite, and the inhibitory input terminated on the soma. The interneurons received excitatory afferent inputs from one or both eyes.

Intracellular recordings from cortical pyramidal neurons *in vivo* typically exhibit large irregular membrane potential fluctuations, which are presumably due to the arrival of synchronous volleys of excitatory postsynaptic potentials[34]. Because our *in vivo* recording data showed similar membrane potential fluctuations, we included background synaptic inputs in our model. Both the proximal and the distal dendrites of the pyramidal neuron received ten excitatory background spike trains, whereas the soma received ten inhibitory background spike trains.

The excitatory and inhibitory inputs were described as non-stationary Poisson spike trains that exhibited spike bursts of duration 50–100 ms and 5–10 ms, respectively, at an average interburst interval of 1 s. The spike frequency in a burst was 10 kHz. All background spike trains were mutually independent. The thalamic afferents were represented by ten independent non-stationary Poisson spike trains of duration 2 s, all of which projected to the distal dendrite of the pyramidal cell and the single-compartment inhibitory neuron. The temporal structure of the afferent spike trains was similar to that of the background spike trains, except that spike bursts arrived at a rate of 3 Hz for the duration of afferent stimuli.

The thalamic afferents and excitatory background inputs were α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA)-receptor-mediated glutamatergic synapses. Interneuron-to-pyramidal-cell synaptic transmission was mediated by GABA$_A$ receptors located on the somatic compartment of the pyramidal neuron. The AMPA and GABA$_A$ synapses were set to their maximum values immediately after a presynaptic spike and otherwise decayed with a time constant of 5 ms. The reversal potentials of the synaptic currents were chosen to be $E_{AMPA} = 0$ mV and $E_{GABA} = -75$ mV. The maximum synaptic conductance was common to all afferent inputs from the contralateral and ipsilateral eye, and was taken independently from $g^{AMPA}_{pyr} = 0.01, 0.02, 0.03, ..., 0.20$ nS cm$^{-2}$ for the pyramidal neuron and from $g^{AMPA}_{int} = 1.0, 3.0, 5.0, ..., 40.0$ nS cm$^{-2}$ for the interneurons. The spike bias response strength was tested for all possible contralateral–ipsilateral and pyramidal–interneuron combinations of these values. The conductances of the remaining synapses were chosen to be $g^{AMPA(BG)}_{pyr} = 0.05$ nS cm$^{-2}$, $g^{GABA(BG)}_{pyr} = 1.0$ nS cm$^{-2}$, ..., $g^{AMPA(BG)}_{int} = 0.2$ μS cm$^{-2}$, $g^{GABA(BG)}_{int} = 1.0$ nS cm$^{-2}$ and $g^{GABA}_{int-pyr} = 1.0$ nS cm$^{-2}$. In the monocular interneurons with modifiable GABAergic synapses, $g^{GABA}_{int-pyr} = 0.5, 0.6, 0.7, ..., 2.4$ nS cm$^{-2}$ and $g^{AMPA}_{intt} = 1.0$ nS cm$^{-2}$.

Our pyramidal-cell model has a spherical soma and cylindrical basal, proximal and distal dendritic compartments, including spike-generating sodium, persistent sodium, delayed-rectifier potassium, slowly inactivating potassium and high-voltage-activated calcium currents. The mathematical details of our model, including the morphological and electrophysiological properties, are identical to those of the pyramidal neuron modelled in ref. 35 for the 0%-dopamine case, except that our model includes no BK-type potassium current.

Our fast-spiking GABAergic interneuron model has a single compartment, and includes spike-generating sodium and Kv3.1 delayed-rectifier potassium currents[36]. The mathematical details have been described previously[37].

31. Coleman, M. J. & Mooney, R. Synaptic transformations underlying highly selective auditory representations of learned birdsong. *J. Neurosci.* **24**, 7251–7265 (2004).
32. Volgushev, M., Pernberg, J. & Eysel, U. T. Comparison of the selectivity of postsynaptic potentials and spike responses in cat visual cortex. *Eur. J. Neurosci.* **12**, 257–263 (2000).
33. Izhikevich, E. M. Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**, 1569–1572 (2003).
34. Stevens, C. F. & Zador, A. M. Input synchrony and the irregular firing of cortical neurons. *Nature Neurosci.* **1**, 210–217 (1998).
35. Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* **83**, 1733–1750 (2000).
36. Erisir, A., Lau, D., Rudy, B. & Leonard, C. S. Function of specific K$^+$ channels in sustained high-frequency firing of fast-spiking neocortical interneurons. *J. Neurophysiol.* **82**, 2476–2489 (1999).
37. Nomura, M., Fukai, T. & Aoyagi, T. Synchrony of fast-spiking interneurons interconnected by GABAergic and electrical synapses. *Neural Comput.* **15**, 2179–2198 (2003).

# LETTERS

# Human *DAZL*, *DAZ* and *BOULE* genes modulate primordial germ-cell and haploid gamete formation

Kehkooi Kee[1], Vanessa T. Angeles[1], Martha Flores[1], Ha Nam Nguyen[1] & Renee A. Reijo Pera[1]

The leading cause of infertility in men and women is quantitative and qualitative defects in human germ-cell (oocyte and sperm) development. Yet, it has not been possible to examine the unique developmental genetics of human germ-cell formation and differentiation owing to inaccessibility of germ cells during fetal development. Although several studies have shown that germ cells can be differentiated from mouse and human embryonic stem cells, human germ cells differentiated in these studies generally did not develop beyond the earliest stages[1–8]. Here we used a germ-cell reporter to quantify and isolate primordial germ cells derived from both male and female human embryonic stem cells. By silencing and overexpressing genes that encode germ-cell-specific cytoplasmic RNA-binding proteins (not transcription factors), we modulated human germ-cell formation and developmental progression. We observed that human *DAZL* (deleted in azoospermia-like) functions in primordial germ-cell formation, whereas closely related genes *DAZ* and *BOULE* (also called *BOLL*) promote later stages of meiosis and development of haploid gametes. These results are significant to the generation of gametes for future basic science and potential clinical applications.

Historically, human germ-cell development has been intractable to direct analysis; yet, infertility is unusually common in both men and women, with genetic requirements that differ from those of other commonly studied species[9,10]. Here, we sought to develop a system for direct experimental examination of landmark events and genetic requirements in human germ-cell formation, maintenance of pluripotency, epigenetic reprogramming and progression through meiosis (Supplementary Fig. 1). Although previous studies had demonstrated that bone morphogenetic proteins (BMPs) promote differentiation of human embryonic stem (ES) cells to germ cells in embryoid bodies, the process was inefficient[5]. Thus, we explored adherent differentiation of human ES cells and observed the induction of a variety of morphological changes (Supplementary Fig. 2). Furthermore, differentiation was accompanied by increased expression of the germ-cell-specific proteins VASA and DAZL, in all human ES cell lines tested (two female (XX) and two male (XY) lines from four independent derivations; Fig. 1a and Supplementary Fig. 3).

On the basis of these data and previous studies indicating that *VASA* is germ-cell specific[4,5,11,12], we constructed a green fluorescent protein (GFP)-conjugated VASA reporter to purify germ cells from the complex cell mixture resulting from human ES cell differentiation (Supplementary Fig. 4). We introduced the reporter into undifferentiated human ES cells, and then following differentiation, isolated GFP+ cells (putative primordial germ cells (PGCs)) via fluorescence-activated cell sorting (FACS) (Fig. 1b). We observed that both XY- and XX-bearing human ES cells reproducibly gave rise to a GFP+ population after 7 and 14 days of differentiation, and that the percentage of GFP+ cells reached approximately 5% with addition of BMPs, which are required for mouse PGC formation[13] (Supplementary Fig. 5). Protein analysis

confirmed that the GFP+ cells are enriched for endogenous VASA and DAZL proteins (Fig. 1c). VASA protein was localized specifically to the cytoplasm of the GFP+ cells and was not detected in GFP− cells (Supplementary Fig. 6a). Further analysis showed that, as expected, OCT4 protein was expressed most highly in undifferentiated human ES cells but also in both GFP+ and GFP− populations at lower levels due to differentiation (Supplementary Fig. 6b).
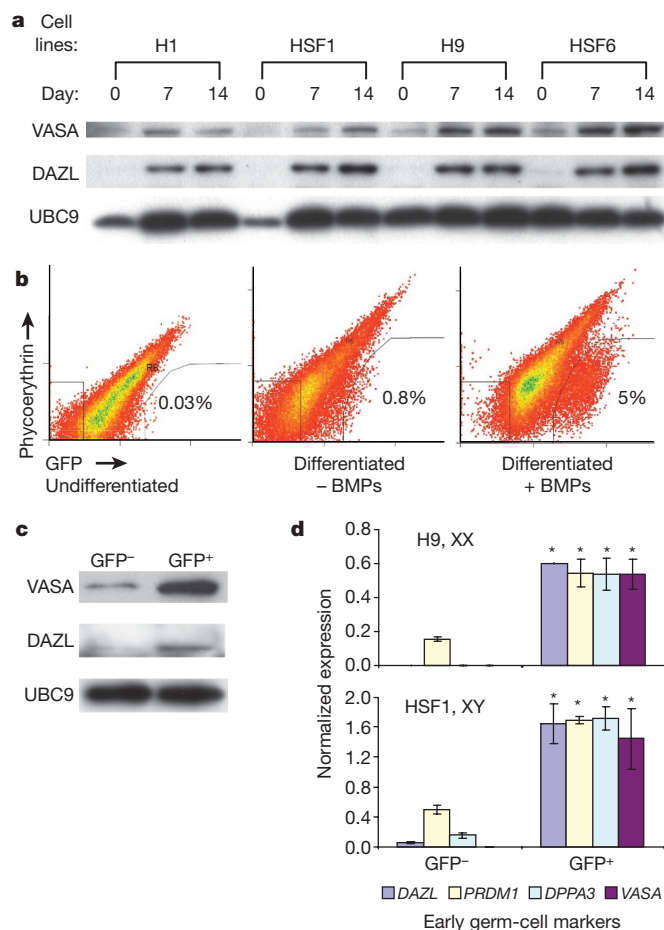


**Figure 1 | Enrichment of human germ cells by BMPs and VASA–GFP reporter. a,** Western blot analysis of human ES cells after differentiation with BMPs at 7 and 14 days. Equal amounts of cell lysates were loaded in each lane. UBC9 was used as a loading control. **b,** FACS analysis of GFP populations. **c,** Western blot analysis of VASA and DAZL after FACS. **d,** Expression of early germ-cell markers via qPCR Taqman probes. A total of 20,000 GFP+ and GFP− cells (day 7 of differentiation with BMPs) were subjected to qPCR analysis. Error bars indicate s.d.; asterisk, significant difference by *t*-test ($P < 0.05$), $n = 2$.

[1]Center for Human Embryonic Stem Cell Research and Education, Institute for Stem Cell Biology & Regenerative Medicine, Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford University, Palo Alto, California 94305, USA.

Gene expression profiling was carried out on the GFP[+] and GFP[−] populations. Early germ-cell markers such as *DAZL*, *PRDM1* (also called *BLIMP1*), *DPPA3* (also called *STELLA*) and *VASA* (also known as *DDX4*) were significantly enriched in the GFP[+] populations (Fig. 1d), whereas those typically expressed during later stages of germ-cell development were either not detected or not enriched, with the exception of low levels of synaptonemal complex protein 3 (*SCP3*) in the GFP[+] population (Supplementary Fig. 7). γH2AX and SCP3 immunostaining was used to examine meiotic progression throughout our experiments; γH2AX is an indicator of meiotic recombination based on binding to double-strand breaks[14,15] and SCP3 is indicative of synaptonemal complex formation in meiotic prophase I[16]. When the cells were stained for SCP3 and γH2AX, we observed that the GFP[+] cells showed only low levels of scattered, punctate SCP3 staining in rare cells and there was no staining of γH2AX (Supplementary Fig. 8). These results indicated that the GFP[+] cells are probably at a pre-meiotic stage (with rare cells entering meiosis). We also observed that GFP[+] cells were enriched for expression of a subset of pluripotency genes—*LIN28*, *NANOG*, *OCT4* (also called *POU5F1*) and *TERT*—consistent with previous reports of their expression in human germ cells[17,18].

Epigenetic reprogramming is diagnostic of germ-cell development[19]. Thus, we characterized erasure of methylation (hypomethylation) globally and at the differentially methylated regions (DMRs) of imprinted loci. We found that the *H19* locus was hypomethylated in GFP[+] cells relative to GFP[−] cells (Fig. 2a). Results from other imprinted loci (*PEG1* (also called *MEST*), *SNRPN*, *KCNQ*) confirmed that the GFP[+] cells also showed significantly lower levels of methylation at these DMRs relative to other cell types (Supplementary Fig. 9). Furthermore, examination of global DNA methylation levels (5-methylcytosine (5mC); Fig. 2b) provided strong evidence that the VASA–GFP[+] population is in the process of erasing methylation globally. Identity of the populations of GFP[−] and GFP[+] cells was verified by staining for VASA and OCT4 (Supplementary Fig. 6).

PGCs possess the ability to establish embryonic germ-cell lines with diagnostic gene expression and morphology[3,20]. Thus, we tested whether GFP[+] cells form embryonic germ lines on inactivated feeder cells in media lacking basic fibroblast growth factor (bFGF). We found that the GFP[+] cells gave rise to colonies that resembled embryonic germ cells[21] after 7 days (Fig. 2c and Supplementary Fig. 10a), whereas the GFP[−] cells did not give rise to any colonies. Similar to embryonic germ cells[22], replated GFP[+] cells had intense alkaline phosphatase activity (Supplementary Fig. 10b) and remained GFP[+] after extensive culture (Fig. 2c). Gene expression profiles of replated cells (after 20 days) were similar to those of freshly isolated GFP[+] cells, with a few exceptions (Supplementary Fig. 11). We noted, however, that DMRs of replated cells had significantly more methylation after replating (Supplementary Fig. 9). This is similar to previous reports with human ES cells[23] but had not been examined in human embryonic germ cells.

Because gene expression, immunostaining, epigenetic status and the ability to give rise to colonies resembling embryonic germ cells strongly suggested that the GFP[+] cells are PGCs, we next examined genetic requirements for formation and differentiation of human PGCs. We focused on the human *DAZ* gene family which contains three members: four human *DAZ* genes which are commonly deleted from the Y chromosome of infertile men who lack germ cells[24,25], and autosomal *DAZL*
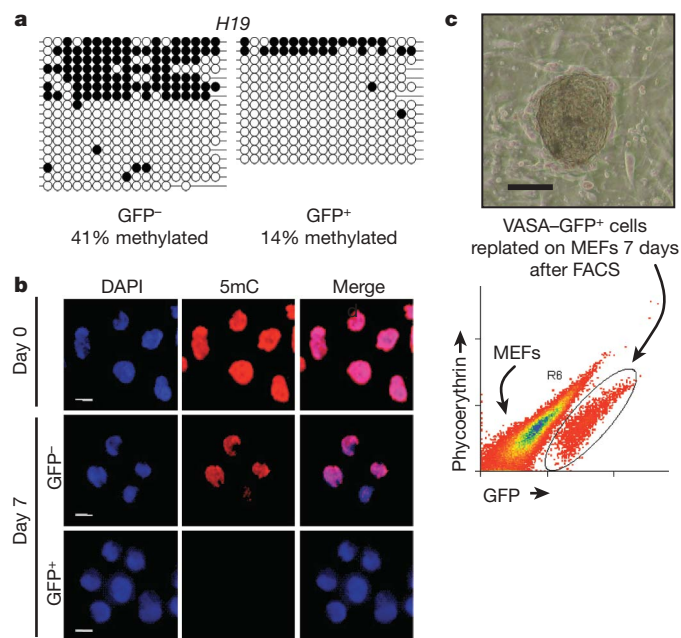
**Figure 3 | Silencing of *DAZ* family members and germ-cell numbers.**
**a**, Western blot analysis of DAZL-V5 silenced by independent *shDAZL* constructs. DAZL with V5 epitope was co-transfected with control (*shLacZ*) or *shDAZL* constructs in 293T cells. Mut2DAZL-V5 was resistant to *shDAZL2*. **b**, Western blot analysis of DAZL and VASA after silencing with *shDAZL4* or rescue in human ES cells (H9). oem4DAZL, Overexpressed DAZL with synonymous mutation resistant to *shDAZL4*. **c**, FACS results using H9 (XX line) and HSF1 (XY line) for silencing of DAZL, BOULE and DAZ. Error bars indicate s.d.; asterisk, significant difference in percentage of VASA–GFP[+] cells by one-way analysis of variance (ANOVA; $P < 0.05$), $n = 2$ (averages from two independently differentiated samples at 14 days).

**Figure 2 | Germ-cell properties of VASA–GFP[+] cells. a**, The GFP[+] population was hypomethylated at the *H19* locus. **b**, 5-methylcytosine staining of the VASA–GFP[+] population to detect global methylation. Cells were immunostained using monoclonal 5-methylcytosine antibody. Images are taken at the same exposure time to show different levels of staining. Scale bars, 10 μm. **c**, Phase contrast pictures showing representative colony from GFP[+] cells after 7 days of replating. No colonies were observed from plating of the GFP[−] population. FACS plot demonstrates that GFP[+] cells maintained GFP expression after 7 days of replating on mouse embryonic fibroblasts (MEFs). Scale bar, 100 μm.
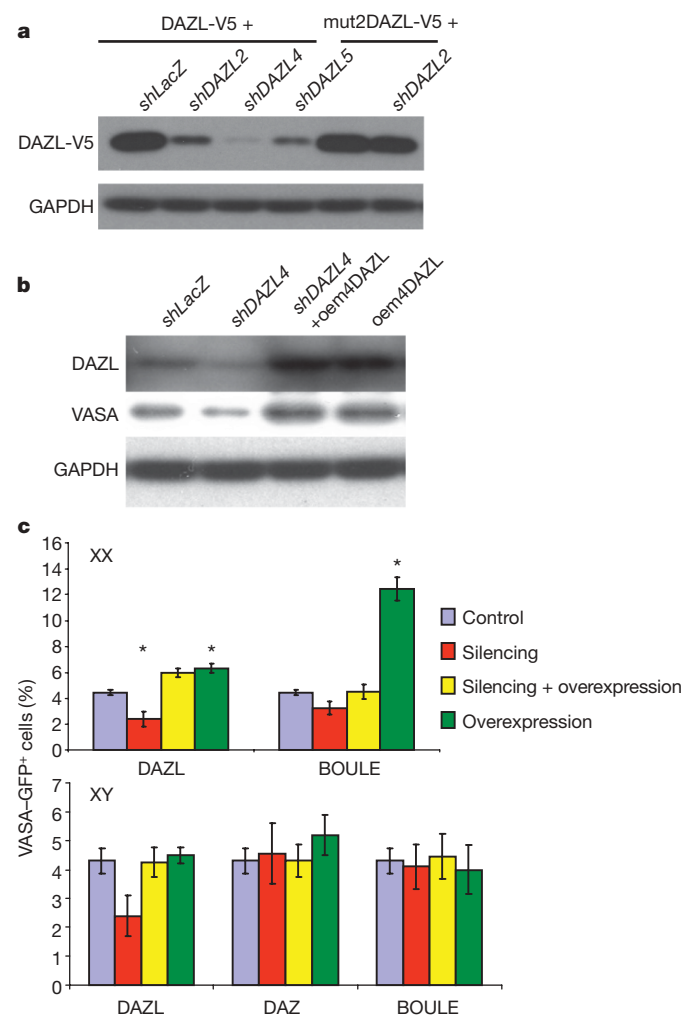
and *BOULE* homologues which are conserved from invertebrates to humans[10]. We first silenced expression of the autosomal *DAZL* gene by short hairpin RNA (shRNA) technology (Supplementary Fig. 12a, b) and observed reduced protein levels with three silencing constructs (Fig. 3a); specificity of silencing was confirmed with synonymous mutations (Supplementary Fig. 12c). Changing just three nucleotides in the *DAZL* shRNA (*shDAZL*) targeting region created a mutated DAZL-V5 resistant to silencing (Fig. 3a). When human ES cells were stably integrated with DAZL silencing vectors and differentiated for 14 days with BMPs, expression of DAZL and VASA was significantly reduced in cells carrying the *shDAZL4* or *shDAZL2* vectors (Fig. 3b and Supplementary Fig. 13a). Overexpression of mutated DAZL resulted in rescue as observed by elevated DAZL and VASA protein expression. Notably, overexpression of DAZL alone elevated endogenous VASA levels relative to controls, indicating that VASA is regulated by DAZL.

To examine whether silencing of BOULE or DAZ affects PGC differentiation, we identified multiple *shBOULE* and *shDAZ* constructs that significantly reduced expression of these proteins (Supplementary Fig. 13b, c). We introduced the silencing constructs into human ES cells carrying the VASA–GFP reporter with individual shRNAs: *shDAZL* and *shBOULE* separately into H9 (XX), and *shDAZL*, *shBOULE* and *shDAZ* into HSF1 (XY). Corresponding target genes were also co-transduced to rescue silencing effects. We observed that the VASA–GFP+ population was reduced to almost half by silencing DAZL in both XX and XY lines (Fig. 3c and Supplementary Fig. 14;

statistically significant in XX and marginal in XY cells). In contrast, silencing of BOULE reduced the GFP+ population slightly in the XX line but not in the XY line, and the number of GFP+ cells was unaffected when DAZ was silenced. Finally, overexpression of BOULE alone increased the VASA–GFP+ population to nearly 12% in XX but not in XY cells, suggesting that BOULE has a more important role in directing human female PGC differentiation than male. Overexpression of combinations of DAZL, BOULE and DAZ did not result in synergistic enhancement of PGC formation (data not shown).

The data above suggested that overexpression of DAZL, and BOULE (in XX lines), promotes PGC formation. We next examined if overexpression of these genes promotes germ-cell differentiation beyond the PGC stage. Thus, combinations of vectors that overexpressed *DAZ* gene family members were introduced into XX and XY human ES cells. Human ES cells were differentiated for 7 days in the absence of BMPs (to test whether internal factors alone can induce late germ-cell differentiation) and examined for meiotic progression. The majority of nuclei from differentiated human ES cells showed no obvious γH2AX staining (including nuclei that did not enter meiosis and those that might have completed meiosis) (Fig. 4a). When γH2AX was detected in the nuclei, it was accompanied by more than ten punctate SCP3 foci, indicating that the nuclei had entered meiotic prophase I. Elongated SCP3 localization with different intensities and lengths was also detected but was not accompanied by γH2AX staining (Fig. 4b). This staining pattern probably corresponds to that of synaptonemal complexes at zygotene, pachytene
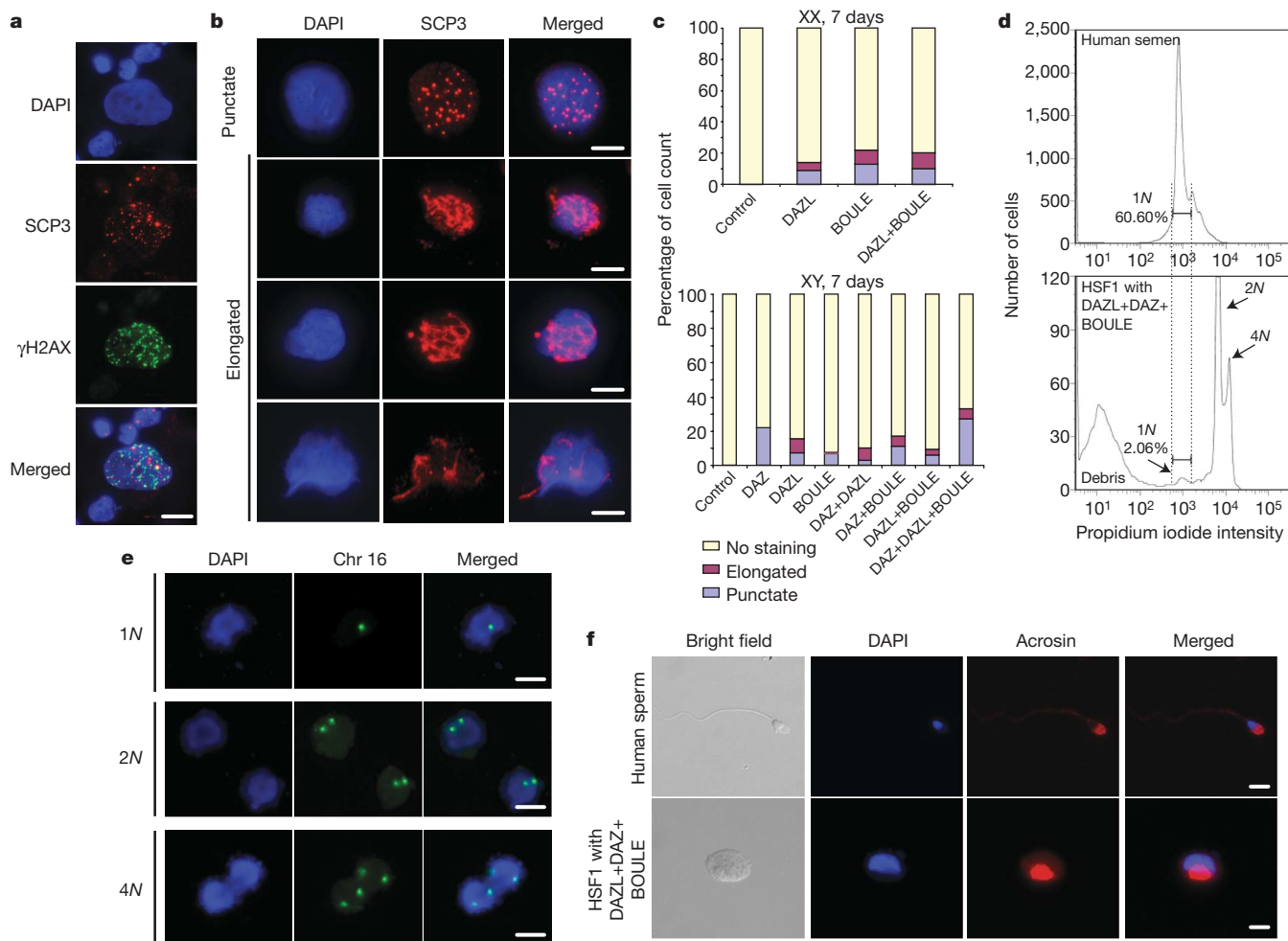


**Figure 4 | Overexpression of DAZL, DAZ and BOULE induces meiotic progression and haploid formation. a,** Meiotic spread from day 7 differentiated cells; immunofluorescence staining with SCP3 and γH2AX is shown. **b,** Meiotic cells overexpressing DAZ family proteins and stained for SCP3. **c,** Percentage of cells showing punctate or elongated SCP3 staining at day 7. A total of 200 meiotic spreads were counted and categorized for each

sample. **d,** FACS of DNA content of human semen and cells overexpressing DAZL, DAZ and BOULE. **e,** Fluorescent *in situ* hybridization of chromosome 16 in cells sorted as 1*N*, 2*N*, 4*N*. **f,** Acrosin staining of 1*N* population from human semen and HSF1 with three overexpression vectors. All cells were from whole cultures without GFP FACS. All scale bars, 10 μm.

or diplotene stages. To quantify synaptonemal complex formation, we categorized SCP3 staining as punctate or elongated, and counted the percentage of nuclei in each category and those negative for SCP3. Overexpression of DAZL, BOULE, or a combination of both gave similar results at day 7 in the XX line (Fig. 4c). In XY cells, overexpression of DAZ, DAZL and BOULE resulted in the highest percentage of cells with SCP3 staining (Fig. 4c and Supplementary Fig. 15). Indeed, overexpression of DAZ alone gave rise to more than 20% of cells with punctate SCP3, much more than DAZL or BOULE alone. We infer that overexpression of DAZL or BOULE was sufficient to induce elongated synaptonemal complex formation in the XY line, but DAZ was required to achieve the highest level of synaptonemal complex formation.

We next determined if haploid cells were produced. We found that messenger RNA expression of the mature sperm markers TEKT1 and acrosin (ACR) was highly elevated in cells that overexpressed all three family members at day 14 (Supplementary Fig. 16), suggesting potential formation of haploid gametes. Moreover, when cells were sorted by DNA content (using parameters developed to sort 1N cells from a human semen sample obtained from the Stanford IVF Clinic), ~2% of cells had 1N content on day 14 after overexpression of DAZL, BOULE and DAZ. No corresponding haploid cells were isolated from control cells that lacked overexpression of DAZ gene family members (Fig. 4d and Supplementary Fig. 17). DNA content was confirmed by fluorescent in situ hybridization (FISH) using a probe for chromosome 16. As expected, sorted 1N cells possessed a single chromosome 16, whereas 2N and 4N cells carried 2 and 4 chromosomes, respectively (Fig. 4e and Supplementary Figs 18a and 19). In addition, most 1N cells also expressed the mature sperm protein ACR, which is present from spermatid to spermatozoan stage[26] (Fig. 4f and Supplementary Figs 18b and 20). In contrast, 2N cells differentiated in the same culture were negative for ACR. We also observed that the H19 DMR was hypermethylated in 1N cells, whereas SNRPN and KCNQ DMRs were hypomethylated with patterns similar to those detected in human semen (Supplementary Fig. 21). Finally, we observed that expression of the genes AMH (also called MIS), FSHR, LHCGR (also called LHR) and SOX9 was greater in cultures that produced the highest number of germ cells (overexpressed DAZ, DAZL and BOULE proteins; in whole culture without FACS), indicating increased numbers of Sertoli and Leydig cells[7] in the same differentiated cultures to support maturation of male germ cells (Supplementary Fig. 22).

Our results indicate that human germ cells can be differentiated and isolated from pluripotent human ES cells and that they possess the ability to enter and progress through meiosis. Moreover, we observed that members of the human DAZ gene family that encode translational regulators modulate both germ-cell formation and differentiation. The human DAZL gene functions primarily in PGC formation, whereas DAZ and BOULE function primarily to promote germ-cell progression to meiosis and formation of haploid germ cells that resemble round spermatids in cellular and molecular characteristics.

## METHODS SUMMARY

**VASA–GFP reporter, transduction and FACS.** 2.5 kb of human VASA upstream of the first codon was cloned into pENTR 5′-TOPO. eGFP was fused 1 kb downstream of the last codon of human VASA, and cloned into pENTR/D-TOPO. Cloned plasmids were recombined[27] to create pLVGV. Lentiviral supernatant was produced, human ES cells were transduced overnight on matrigel in conditioned medium and subsequently selected with geneticin (200 ng ml$^{-1}$) for 7 days. Selected human ES cells were differentiated for the times indicated and harvested by brief treatment with collagenase IV and then TrypLEExpress (Invitrogen). The cell suspension was prepared in differentiation medium for FACS with a MoFlow or BD cell sorter.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 June; accepted 12 October 2009.
Published online 28 October; corrected 12 November 2009 (see full-text HTML version for details).

1. Hubner, K. et al. Derivation of oocytes from mouse embryonic stem cells. Science 300, 1251–1256 (2003).
2. Toyooka, Y., Tsunekawa, N., Akasu, R. & Noce, T. Embryonic stem cells can form germ cells in vitro. Proc. Natl Acad. Sci. USA 100, 11457–11462 (2003).
3. Geijsen, N. et al. Derivation of embryonic germ cells and male gametes from embryonic stem cells. Nature 427, 148–154 (2004).
4. Clark, A. T. et al. Spontaneous differentiation of germ cells from human embryonic stem cells in vitro. Hum. Mol. Genet. 13, 727–739 (2004).
5. Kee, K., Gonsalves, J. M., Clark, A. T. & Reijo Pera, R. A. Bone morphogenetic proteins induce germ cell differentiation from human embryonic stem cells. Stem Cells Dev. 15, 831–837 (2006).
6. Tilgner, K. et al. Isolation of primoridal germ cells from differentiating human embryonic stem cells. Stem Cells 26, 3075–3085 (2008).
7. Bucay, N. et al. A novel approach for the derivation of putative primordial germ cells and Sertoli cells from human embryonic stem cells. Stem Cells 27, 68–77 (2009).
8. Park, T. S. et al. Derivation of primordial germ cells from human embryonic and induced pluripotent stem cells is significantly improved by coculture with human fetal gonadal cells. Stem Cells 27, 783–795 (2009).
9. Hull, M. G. R. et al. Population study of causes, treatment, and outcome of infertility. Br. Med. J. (Clin. Res. Ed.) 291, 1693–1697 (1985).
10. Xu, E. Y., Moore, F. L. & Reijo Pera, R. A. A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. Proc. Natl Acad. Sci. USA 98, 7414–7419 (2001).
11. Fujiwara, Y. et al. Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosophila vasa and its specific expression in germ cell lineage. Proc. Natl Acad. Sci. USA 91, 12258–12262 (1994).
12. Castrillon, D. H., Quade, B. J., Wang, T. Y., Quigley, C. & Crum, C. P. The human VASA gene is specifically expressed in the germ cell lineage. Proc. Natl Acad. Sci. USA 97, 9585–9590 (2000).
13. Zhao, G. Q. Consequences of knocking out BMP signaling in the mouse. Genesis 35, 43–56 (2003).
14. Mahadevaiah, S. K. et al. Recombinational DNA double-strand breaks in mice precede synapsis. Nature Genet. 27, 271–276 (2001).
15. Lenzi, M. L. et al. Extreme heterogeneity in the molecular events leading to the establishment of chiasmata during meiosis I in human oocytes. Am. J. Hum. Genet. 76, 112–127 (2005).
16. Heyting, C. Synaptonemal complexes: structure and function. Curr. Opin. Cell Biol. 8, 389–396 (1996).
17. Perrett, R. M. et al. The early human germ cell lineage does not express SOX2 during in vivo development or upon in vitro culture. Biol. Reprod. 78, 852–858 (2008).
18. Miyamoto, T. et al. Isolation and expression analysis of the testis-specific gene, STRA8, stimulated by retinoic acid gene 8. J. Assist. Reprod. Genet. 19, 531–535 (2002).
19. Hajkova, P. et al. Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. 117, 15–23 (2002).
20. West, J. A. et al. A role for Lin28 in primordial germ-cell development and germ-cell malignancy. Nature 460, 909–913 (2009).
21. Shamblott, M. J. et al. Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc. Natl Acad. Sci. USA 95, 13726–13731 (1998).
22. McLaren, A. Primordial germ cells in the mouse. Dev. Biol. 262, 1–15 (2003).
23. Rugg-Gunn, P. J., Ferguson-Smith, A. C. & Pedersen, R. A. Epigenetic status of human embryonic stem cells. Nature Genet. 37, 585–587 (2005).
24. Reijo, R. et al. Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. Nature Genet. 10, 383–393 (1995).
25. Reijo, R., Alagappan, R. K., Patrizio, P. & Page, D. C. Severe oligospermia resulting from deletions of the Azoospermia Factor gene on the Y chromosome. Lancet 347, 1290–1293 (1996).
26. Flörke-Gerloff, S., Topfer-Petersen, E., Muller-Esterl, W., Schill, W. B. & Engel, W. Acrosin and the acrosome in human spermatogenesis. Hum. Genet. 65, 61–67 (1983).
27. Suter, D. M. et al. Rapid generation of stable transgenic embryonic stem cell lines using modular lentivectors. Stem Cells 24, 615–623 (2006).

# METHODS

**Human ES cell lines and adherent differentiation.** Four human ES cell lines were used in this study: HSF1(XY), HSF6(XX), H1(XY) and H9(XX). Undifferentiated cultures of human ES cells were maintained on irradiated MEFs as previously described[5]. Briefly, all cultures were grown at 37 °C with 5% $CO_2$ in knockout serum replacer (KSR) plus bFGF medium (knockout DMEM, supplemented with 20% knockout serum replacer, 1 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol, and 4 ng ml$^{-1}$ recombinant human bFGF (bFGF, R&D systems)). To adherently differentiate human ES cells on matrigel, about $5 \times 10^4$ human ES cells (less than 50% confluency of a six-well plate) were treated with collagenase type IV (1 mg ml$^{-1}$) for 10 min, scraped with a 5 ml plastic pipette, and transferred with fresh KSR + bFGF media to an identical well coated with matrigel. Conditioned medium (KSR + bFGF media collected after overnight incubation on irradiated MEFs) was used to maintain the undifferentiated cells whenever drug selection (geneticin at 200 µg ml$^{-1}$, and zeocin and blasticidin at 2 µg ml$^{-1}$) was required. Differentiation began upon aspiration of KSR + bFGF media or conditioned media, washed once with PBS without $Ca^{2+}$ and $Mg^{2+}$, and replaced with differentiation media (knockout DMEM supplemented with 20% fetal bovine serum, 1 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol and 50 ng ml$^{-1}$ recombinant human BMP4, BMP7 and BMP8b (R&D systems)). Differentiation medium was changed after 7 days of incubation if longer differentiation was necessary.

**Western blot analysis.** To collect the adherently differentiated cells, first they were washed with 3 ml cold PBS without $Ca^{2+}$ and $Mg^{2+}$, then scraped from the plate in 1 ml PBS plus 2× protease inhibitors (Complete Mini, Roche) and transferred immediately to a 1.5-ml tube on ice. The PBS cell suspension was then spun at 5,000 r.p.m. in a microcentrifuge for 3 min and the supernatant was discarded. The cell pellet was resuspended with 200 µl RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% sodium deoxycholate, 1% NP-40, 0.1% SDS, pH 8) plus 2× protease inhibitors (Complete Mini, Roche). The cell pellet suspension was pipetted rigorously at least 10 times, then vortexed for 30 s. The suspension was again spun down for 3 min at the same speed. The supernatant was measured for protein concentration and denatured at a 1:1 ratio with 2× Laemmli buffer at 95 °C for 5 min, then loaded on to either a 10% or 12% SDS–PAGE gel. The SDS–PAGE gels were run at 150 V for 1 h and transferred to a PVDF membrane for 1 h at 100 V in CAPS buffer (10 mM CAPS, 10% methanol, pH 11). Transferred blots were blocked in 5% non-fat milk for 1 h at room temperature. The blot was subjected to 1 h of primary antibody incubation (1:500 for anti-VASA (Abcam), 1:500 anti-DAZL-150[28], 1:1,000 for anti-V5 (Abcam), 1:10,000 for anti-UBC9 (Abcam) and anti-GAPDH (Abcam)), followed by two quick rinses and three washes for 5 min in TBST (TBS, pH 7.4 with 0.1% Tween-20). Secondary antibody (1:10,000 anti-rabbit-HRP conjugated (Amersham)) incubation had the same duration and washes. ECL+ (Amersham) was used to detect the HRP signal on film.

**Quantitative PCR on mRNA expression and statistical analysis.** Total RNA was collected using the RNeasy system (Qiagen) or PicoPure RNA isolation (Arcturus) and cDNA was prepared with SuperScript III Reverse Transcriptase (Invitrogen) according to manufacturer's protocols. Quantitative PCR was conducted with Taqman probes (Applied Biosystems) using the Fluidigm System (Biomark). Gene expression is normalized to a set of housekeeping genes (*GAPDH*, *RPLP0*, centrin, *CTNNB1*, *ACTB*) and calculated according to geNorm program[29] for GFP$^-$ versus GFP$^+$ cells. Gene expressions of the replated cells, fetal liver cDNA (Clontech), and overexpressed cells were normalized again to either the expression of GFP$^+$ cells or control cells carrying empty p2k7, and the normalized expression value is therefore fold of change. Statistical analysis was calculated using Student's *t*-test or one-way ANOVA by Prism 5.0 program.

**DNA methylation analysis.** VASA–GFP$^+$, VASA–GFP$^-$ and 1*N* populations were collected by FACS; adult skin fibroblast, H9 and HSF1 were collected from trypsinized cell suspensions; replated GFP$^+$ cells were collected by colony picking under dissection microscope (pooled of five colonies). Genomic DNA was extracted using the ZR Genomic DNA II kit (Zymo Research) or QIAamp DNA Mini kit (Qiagen). For bisulphite sequencing analysis of *H19* locus, 100 ng of genomic DNA was processed using the Qiagen Epitect Bisulfite Kit (Qiagen) according to protocol. One microlitre of bisulphite-treated genomic DNA was PCR amplified at the *H19* locus using primers as previously described[4]. The resultant product was gel-extracted using the Qiaquick gel extraction kit (Qiagen) and cloned into the TOPO TA vector (Invitrogen). At least 20 clones were sequenced using ABI BigDye v3.1 dye terminator sequencing chemistry (Applied Biosystems) and ABI PRISM 3730xl capillary DNA analyser for sequence analysis. Quantitative DNA methylation analysis using methylation-sensitive/dependent restriction assay was carried out according to a previous study[5]. Briefly, ~50–200 ng of each genomic DNA sample was divided equally for control (with buffers), methylation-sensitive (NotI, HhaI and HpaII; New England Biolabs) and methylation-dependent (McrBC; New England Biolabs)

restriction digestions. 1/30 of the digested genomic DNA was then used as the input for qPCR reaction (ABI Power SYB master mix) using specific primers for the DMRs of *H19*[30], *PEG1*[30], *SNRPN*[23] and *KCNQ*[23] respectively, and triplicates were ran for each reaction. Percentage of hypermethylation was calculated as previously reported[31], and the average percentages with lower standard deviation were chosen between methylation-sensitive and methylation-dependent restriction reactions.

**Alkaline phosphatase assay.** Cells grown on MEFs were subjected to alkaline phosphatase staining using Vector Red Alkaline Phosphatase Substrate Kit I (Vector Laboratories) according to manufacturer's protocols.

**shRNA vectors and overexpression vectors.** All shRNAs targeting human *DAZL*, *DAZ* and *BOULE* were constructed using the Block-iT inducible H1 lentiviral system (Invitrogen). All shRNAs were first introduced into pENTR/H1/TO vectors and transferred into pLenti4/BLOCK-iT-DEST destination vectors (see below for sequence information). When shRNA was tested in 293FT cells for its efficacy of silencing, transient transfection without selection was carried out and cell lysate was collected after 24 h. After specific shRNA was chosen, viral supernatant was prepared as described for pLVGV. Subsequently, the viral supernatant was used to transduce human ES cells on matrigel as described above, except with zeocin selection (2 µg ml$^{-1}$) for 2 days for stable integration of the shRNA vector into human ES cells. Overexpression vectors were constructed by inserting the EF1α promoter, and *DAZL*, *DAZ2* or *BOULE* cDNA into the p2k7$_{blas}$ vectors, selected with blasticidin (2 µg ml$^{-1}$) in culture for 3 days to ensure stable integration and subjected for each differentiation experiment. shRNA sequences were: *shDAZL1*, GGATGGATGAAACTGAGATTA; *shDAZL2*, GCATATCCTACTTACCCAAAT; *shDAZL3*, GCCAAATGAATGTTCAGTTCA; *shDAZL4*, GCAGAAGATAGTAGAATCACA; *shDAZL5*, GGATATCAGTTGCCTGTATAT; *shDAZL6*, GGTGGTATCTTGTCTGTTTAA; *shDAZL7*, GGTATCTTGTCTGTTTAATCC; *shDAZ1*, GCAAATCCTGAGACTCCAAAC; *shDAZ2*, GGAAGCTGCTTTGGTAGATAC; *shDAZ3*, GCCACGTCCTTTGGTAGTTAA; *shDAZ4*, GCATTTCCTGCTTATCCAAAT; *shDAZ5*, GCATTTCCTGCTTATCCAAGT; *shBOULE1*, GCATCTTTGTAGGAGGAATTG; *shBOULE2*, GGATCCCTCGTTCTAGTATAA; *shBOULE3*, GCTGGAACAATGTATCTAACA; *shBOULE4*, GCAACCTTCTGAGGTTATTTA; *shBOULE5*, GCTCCAAGTGCCATCACTATG.

**VASA and OCT4 whole-cell staining and 5-methylcytosine nuclei staining.** Cell suspensions from FACS were collected onto a glass slide with a Cytospin (Thermo Scientific) at 800 r.p.m. for 5 min followed by a wash in PBS, and fixed with 4% paraformaldehyde/PBS for 15 min. For VASA and OCT4 staining, slides were blocked in 1% BSA/0.1% Tween-20/PBS for 1 h, followed by overnight incubation with primary antibody (1:500 for anti-VASA (Abcam), 1:200 anti-OCT4 (Santa Cruz)). Slides were then rinsed twice and washed three times, 5 min with 0.1% Tween-20/PBS (PBST). Secondary antibody, anti-rabbit-594 (Invitrogen), was incubated for 1 h at room temperature on the slides, followed by the same washes. The slides were then mounted with Prolong Antifade with DAPI (Molecular Probes). For 5mC staining, fixed cells were incubated with 1% Triton X-100 for 1 h, washed with PBST twice for 5 min. Denaturation of DNA was carried out by treating cells with 4 M HCl/0.1% Triton X-100 for 10 min and immediately neutralized with 100 mM Tris/HCl (pH 8.5) for 30 min. Slides were blocked in 1% BSA/0.1% Tween-20/PBS at 4 °C overnight. This was then stained with antibody against 5-methylcytosine (Genway) at a concentration of 1:100 in blocking buffer for 4 h at room temperature. The slides were washed, incubated with secondary antibody, and mounted the same as described above.

**Meiotic spreads, SCP3 and γH2AX staining.** Meiotic spreads were carried out as described previously[32] with some modifications. Briefly, differentiated human ES cells were prepared as single cells as for FACS analysis. Cells were lysed by a hypotonic solution and dropped on glass slides freshly submerged in 1% PFA and 0.15% Triton X-100. Slides were incubated overnight at 4 °C until the suspension was semi-dry. Slides were treated with 0.04% photoflo for 5 min and blocked with 4% goat serum. Rabbit anti-SCP3 (1:1,000, Novus) and mouse-anti-γH2AX (1:200, Abcam) were incubated for 3 h at room temperature. Washes, secondary antibody incubation (1:1,000, anti-rabbit 594 and anti-mouse 488 (Molecular Probe)) and mounting are the same as described above.

**FACS analysis for DNA content.** Single cell suspensions were prepared as described above. Fixing and staining of the DNA was conducted as described previously[33]. Briefly, cells were fixed in 70% ethanol for 1 h at room temperature, followed by incubation in 0.5 ml staining solution (0.1% Triton X-100 in PBS, 0.2 mg ml$^{-1}$ RNaseA, 0.02 mg ml$^{-1}$ propidium iodide) then transferred to a FACS tube and incubated at 37 °C for >15 min. The cell suspension in the staining solution was subjected to FACS analysis.

**FISH of FACS cells.** 1*N*, 2*N* and 4*N* cells were collected by cytospin as described above after the DNA content FACS. Slides were fixed with Carnoy's fixative (1:3 of acetic acid:methanol), 5 min, air dried in chemical hood, followed by dehydration in ice-cold 70%, 80%, 100% ethanol, 2 min, and air dried for 5 min. FISH

probe (against chromosome 16, Vysis) was directly denatured on slides at 85 °C and hybridized at 37 °C overnight. Slides were washed with 2× SSC once, followed by 0.1% SDS in 2× SSC (pre-warm) at 50 °C, 5 min, and mounted as described earlier.

**Acrosin staining.** The FACS-collected 1*N* population was cytospun onto glass slides, washed once with PBS, and fixed with 4% paraformaldehyde/PBS for 15 min. Fixed cells were then incubated with 1% Triton X-100 for 15 min, washed twice with 0.1% Tween-20/PBS for 5 min. Slides were blocked with 4% goat serum for 1 h at room temperature, then stained with antibody against acrosin (Santa Cruz) diluted 1:50 in blocking buffer for 4 h at room temperature. Slides were rinsed twice and washed three times with PBST, for 5 min. The secondary antibody used was anti-rabbit-Cy5 (Molecular Probes) at 1:1,000 for 1 h at room temperature, followed by the same washes, mounting as described above.

**Cloning and silencing of DAZ1–4.** cDNA clones of DAZ1–4 (NCBI accession numbers: DAZ1, BC114927; DAZ2, NM_020363; DAZ3, BC113005; DAZ4, BC047480) were obtained from ATCC, PCR amplified, cloned into pENTR-D-TOPO vector, and subsequently cloned into pLenti4/TO/V5-DEST (Invitrogen). Expression of these cDNAs resulted in ~42 kDa, 63 kDa, 50 kDa and 44 kDa proteins, respectively, on western blot detected by mouse anti-DAZ1 antibody (Abcam). Co-transfection of the overexpression vector and *shDAZ4* at the ratio of 1:1 and 1:2 into 293FT cells was carried out and all the lysates were collected after 24 h and loaded equally onto 10% SDS–PAGE gel.

**Replating of VASA–GFP⁺ cells.** VASA–GFP⁻ and VASA–GFP⁺ cells were collected in ice-cold KSR media without bFGF (knockout DMEM, supplemented with 20% knockout serum replacer, 1 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol). Immediately after FACS collection, cells were distributed into wells plated with inactivated MEFs containing KSR media. Approximately 20,000 cells were distributed into each well of a six-well plate. Culture was then grown at 37 °C with 5% $CO_2$. Media was changed every 3 days. Colonies usually appeared after 7 days and at 1/5,000 survival rate. Colonies expanded much slower than human ES cell colonies and harvested when colony size was about 200 μm for various assays.

28. Reijo, R. A. *et al.* DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice. *Biol. Reprod.* **63,** 1490–1496 (2000).

29. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3,** RESEARCH0034 (2002).

30. Kerjean, A. *et al.* Establishment of the paternal methylation imprint of the human H19 and MEST/PEG1 genes during spermatogenesis. *Hum. Mol. Genet.* **9,** 2183–2187 (2000).

31. Oakes, C. C., La Salle, S., Robaire, B. & Trasler, J. M. Evaluation of a quantitative DNA methylation analysis technique using methylation-sensitive/dependent restriction enzymes and real-time PCR. *Epigenetics* **1,** 146–152 (2006).

32. Gonsalves, J. *et al.* Defective recombination in infertile men. *Hum. Mol. Genet.* **13,** 2875–2883 (2004).

33. Darzynkiewicz, Z. & Juan, G. DNA content measurement for DNA ploidy and cell cycle analysis. *Curr. Protoc. Cytom.* Ch. 7, Unit 7 5 (2001).

# LETTERS

# Uptake through glycoprotein 2 of FimH+ bacteria by M cells initiates mucosal immune response

Koji Hase[1]*, Kazuya Kawano[1]*, Tomonori Nochi[2], Gemilson Soares Pontes[2], Shinji Fukuda[1,3], Masashi Ebisawa[1,3], Kazunori Kadokura[1,3], Toru Tobe[4], Yumiko Fujimura[1], Sayaka Kawano[1], Atsuko Yabashi[5], Satoshi Waguri[5], Gaku Nakato[1,3], Shunsuke Kimura[1], Takaya Murakami[1], Mitsutoshi Iimura[6], Kimiyo Hamura[6], Shin-Ichi Fukuoka[7], Anson W. Lowe[8], Kikuji Itoh[9], Hiroshi Kiyono[2] & Hiroshi Ohno[1,3]

The mucosal immune system forms the largest part of the entire immune system, containing about three-quarters of all lymphocytes and producing grams of secretory IgA daily to protect the mucosal surface from pathogens[1–3]. To evoke the mucosal immune response, antigens on the mucosal surface must be transported across the epithelial barrier into organized lymphoid structures such as Peyer's patches[4]. This function, called antigen transcytosis, is mediated by specialized epithelial M cells[5,6]. The molecular mechanisms promoting this antigen uptake, however, are largely unknown. Here we report that glycoprotein 2 (GP2), specifically expressed on the apical plasma membrane of M cells among enterocytes, serves as a transcytotic receptor for mucosal antigens. Recombinant GP2 protein selectively bound a subset of commensal and pathogenic enterobacteria, including *Escherichia coli* and *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium), by recognizing FimH, a component of type I pili on the bacterial outer membrane. Consistently, these bacteria were colocalized with endogenous GP2 on the apical plasma membrane as well as in cytoplasmic vesicles in M cells. Moreover, deficiency of bacterial FimH or host GP2 led to defects in transcytosis of type-I-piliated bacteria through M cells, resulting in an attenuation of antigen-specific immune responses in Peyer's patches. GP2 is therefore a previously unrecognized transcytotic receptor on M cells for type-I-piliated bacteria and is a prerequisite for the mucosal immune response to these bacteria. Given that M cells are considered a promising target for oral vaccination against various infectious diseases[7,8], the GP2-dependent transcytotic pathway could provide a new target for the development of M-cell-targeted mucosal vaccines.

M cells are atypical epithelial cells that actively phagocytose macromolecules and microbes. Unlike dendritic cells (DCs) and macrophages, which target antigens to lysosomes for degradation, M cells mainly transcytose the internalized antigens[6]. This vigorous macromolecular transcytosis through M cells delivers antigen to the underlying organized lymphoid follicles and is believed to be essential for initiating antigen-specific mucosal immune responses[9]. Previous studies of mucosal immunity have focused mainly on the roles of mucosal DCs and T lymphocytes, with far less attention on M cells. For example, it is unknown whether antigen transcytosis by M cells is non-specific or mediated by specific receptors that bind to antigens with certain molecular patterns[10]. Efficient transport of luminal microorganisms by M cells led us to speculate that M cells

may express transcytotic receptors for particular microorganisms on their apical surface. To identify such receptors we took advantage of a microarray-based approach for a genome-wide survey of M-cell-specific molecules[11,12]. One such molecule, GP2, was previously thought to be expressed exclusively by pancreatic acinar cells of most mammals[13], but our microarray analysis showed that *Gp2* mRNA is also highly expressed in follicle-associated epithelium (FAE) of mouse Peyer's patches, in contrast with intestinal villus epithelium[11]. GP2 expression in FAE was also confirmed by quantitative polymerase chain reaction (qPCR) (Fig. 1a) and immunoblot analyses (Fig. 1b). Data from *in situ* hybridization (Fig. 1c, left) and immunostaining (Fig. 1d) demonstrated that GP2 expression was restricted to M cells in FAE, as depicted by co-staining with an M-cell-recognizing lectin, *Ulex europaeus* agglutinin-1 (UEA-1) (Supplementary Fig. 2; note that UEA-1 also stains goblet cells as indicated by the arrowhead). Apart from M cells from Peyer's patches, GP2 was also widely distributed on M cells in other organized mucosal lymphoid tissues, namely isolated lymphoid follicles, caecal patches and colonic patches, in mice (Fig. 1e). GP2 was also expressed on M cells of human Peyer's patches, as depicted by a similar expression pattern to that of a human M-cell marker, clusterin[14] (Fig. 1c, f, and Supplementary Fig. 3). These observations define GP2 as the first universal M-cell marker in mouse and human.

We next assessed subcellular localization of GP2 in M cells by immunoelectron microscopy. Positive signals for GP2 were localized intensely on the M-cell apical surface (Fig. 1g and Supplementary Fig. 4). This was further confirmed by apical GP2 staining when membranes were not permeabilized (Fig. 1h). GP2 was also detected in the subapical tubulovesicular compartment, most probably endosomes (Fig. 1g, arrowheads), raising the possibility that GP2 could function as an endocytic receptor. To evaluate this possibility, we injected anti-GP2 monoclonal antibody (mAb) into a ligated intestinal loop containing Peyer's patches. Some of the injected mAb remained on the apical plasma membrane of the M cells, but it was also internalized into cytoplasmic vesicular structures (Supplementary Movie 1), supporting a potential function of GP2 on M cells as an endocytic receptor for luminal antigens.

We next sought to identify the luminal antigens bound to GP2. GP2 has a similar domain structure and 52% overall amino acid sequence identity to Tamm–Horsfall protein (THP), which is expressed on the apical plasma membrane of renal tubular epithelium[15]. THP specifically
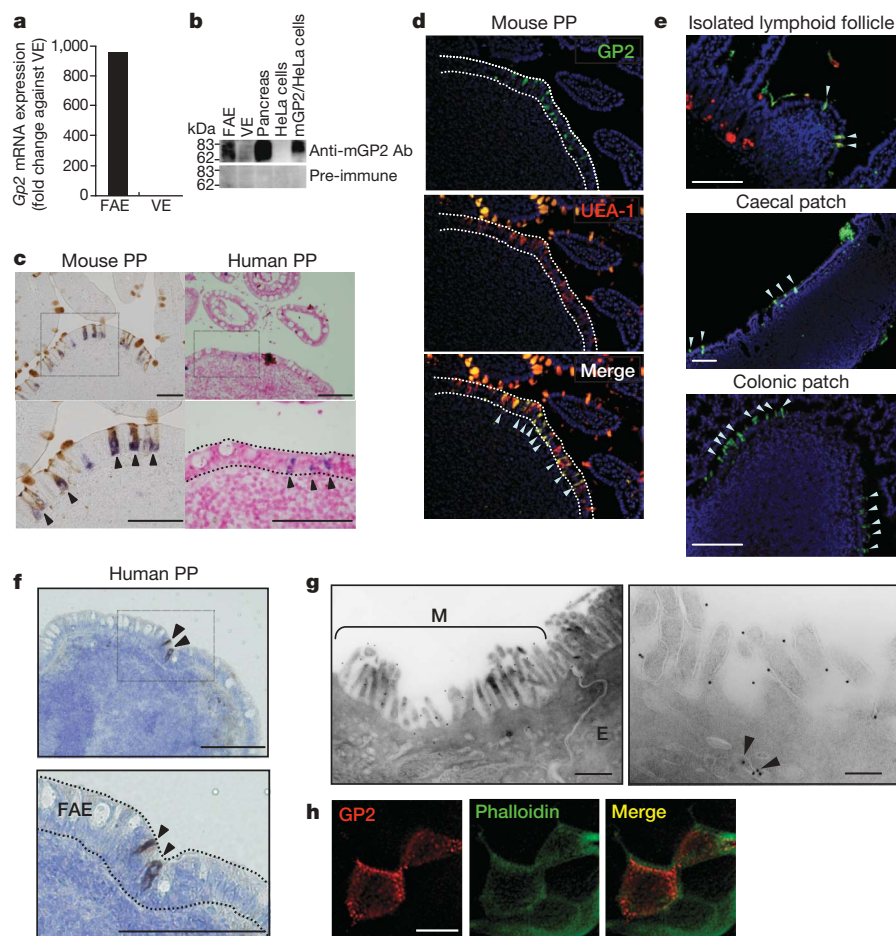
**Figure 1 | In the intestine, GP2 is exclusively expressed by mouse and human M cells. a**, qPCR analysis for mouse *Gp2* mRNA. VE, villus epithelium. **b**, Immunoblotting of mouse GP2 in FAE and villus epithelium. Pancreas and HeLa cells, with or without ectopic expression of a mouse GP2 complementary DNA, are shown as positive controls. Ab, antibody. **c**, *In situ* hybridization analysis using a specific probe for *Gp2* mRNA (purple) in mouse and human Peyer's patch (PP). M cells are co-stained with UEA-1 lectin (brown) in mouse Peyer's patch. Lower panels are magnified images of the areas surrounded by a dotted square in the upper panels. **d**, **e**, Tissue sections prepared from mouse Peyer's patch (**d**) and from isolated lymphoid follicle, caecal patch and colonic patch (**e**) were stained with anti-mouse GP2 mAb (green), UEA-1 (red) and 4,6-diamidino-2-phenylindole (blue). Dotted lines indicate the location of the FAE in the Peyer's patch. GP2-expressing cells scattered in the FAE are indicated by arrowheads. **f**, Immunohistochemical staining of a tissue section of human Peyer's patch

with anti-human GP2 mAb. The positive signal appears brown. The blue colour shows counterstaining with haematoxylin. Dotted lines in the lower panel indicate the location of the FAE in human Peyer's patch. **g**, Subcellular localization of GP2 analysed by immunoelectron microscopy. The M cell (indicated by M) is characterized by sparse and irregular microvilli instead of the rigid and closely packed microvilli observed in normal absorptive enterocytes (indicated by E). GP2 expression (indicated by colloidal gold particles) is visible on the apical plasma membrane and also in the subapical tubulovesicular structures (arrowheads). Scale bars, 100 μm (**c**–**f**), 0.5 μm (**g**, left), 0.2 μm (**g**, right). **h**, The luminal side of whole-mount specimens of a Peyer's patch was treated with anti-GP2 mAb and subsequently with Alexa594-labelled secondary antibody (red) without membrane permeabilization, and then counterstained with phalloidin (green) with membrane permeabilization by 0.1% Triton X-100 in PBS. Scale bar, 10 μm.

binds to uropathogenic *E. coli* and promotes bacterial clearance[16]. We therefore reasoned that GP2 might also bind particular enterobacteria. Indeed, fusion proteins of the Fc fragment of human immunoglobulin G1 with mouse and human GP2 (mGP2-Fc and hGP2-Fc, respectively) both bound *E. coli* (Fig. 2a). The binding of GP2 to *E. coli* is largely mediated by FimH present on the type I pilus tip, because deficiency of FimH abrogated the binding (Fig. 2b). FimH is well documented to bind certain glycoproteins specifically in a mannose-dependent fashion[17]. Consistently, the presence of mannose inhibited the interaction between GP2 and *E. coli* (Supplementary Fig. 5). In sharp contrast to the case in *FimH* deletion, the binding was unaffected in the absence of *CsgA* or *FliC*, which encode the main structural component of curli and flagella, respectively. Lipopolysaccharide and peptidoglycan, non-filamentous components of the bacterial outer membrane, did not bind to GP2 (Supplementary Fig. 6). Thus the FimH⁺ type I pilus binds GP2, which is consistent with a recent study[18]. This view is also supported by our observation that GP2 specifically bound FimH⁺ (*E. coli*, *S. enterica* serovar Enteritidis and *S. Typhimurium*) bacteria[19] but not FimH⁻

(*Pseudomonas aeruginosa*, *Listeria monocytogenes* and *Yersinia enterocolitica*) bacteria (Fig. 2c).

The interaction of GP2 with FimH suggested a possible role of GP2 as a transcytotic receptor for type-I-piliated bacteria. We examined this possibility by the ligated intestinal loop assay using *E. coli* expressing green fluorescent protein (GFP). Three-dimensional imaging analysis of M cells with deconvolution microscopy[20,21] demonstrated the accumulation of GP2 around GFP-*E. coli* just being internalized into M cells through the apical plasma membrane (Fig. 2d, e, and Supplementary Movie 2). Similarly, GP2 was colocalized with GFP-expressing *S. Typhimurium* on the apical surface as well as in subapical vesicular structures (Supplementary Fig. 7 and Movie 3). The number of *FimH*-deficient mutants internalized into M cells was significantly lower than that of the wild-type strain (Supplementary Fig. 8). Furthermore, the uptake of type-I-piliated bacteria by M cells was markedly lower in GP2-deficient (*Gp2*⁻ᐟ⁻) mice than in wild-type (*Gp2*⁺ᐟ⁺) mice (Supplementary Fig. 9 and Fig. 3a, left). In contrast, there was no significant difference in translocation of
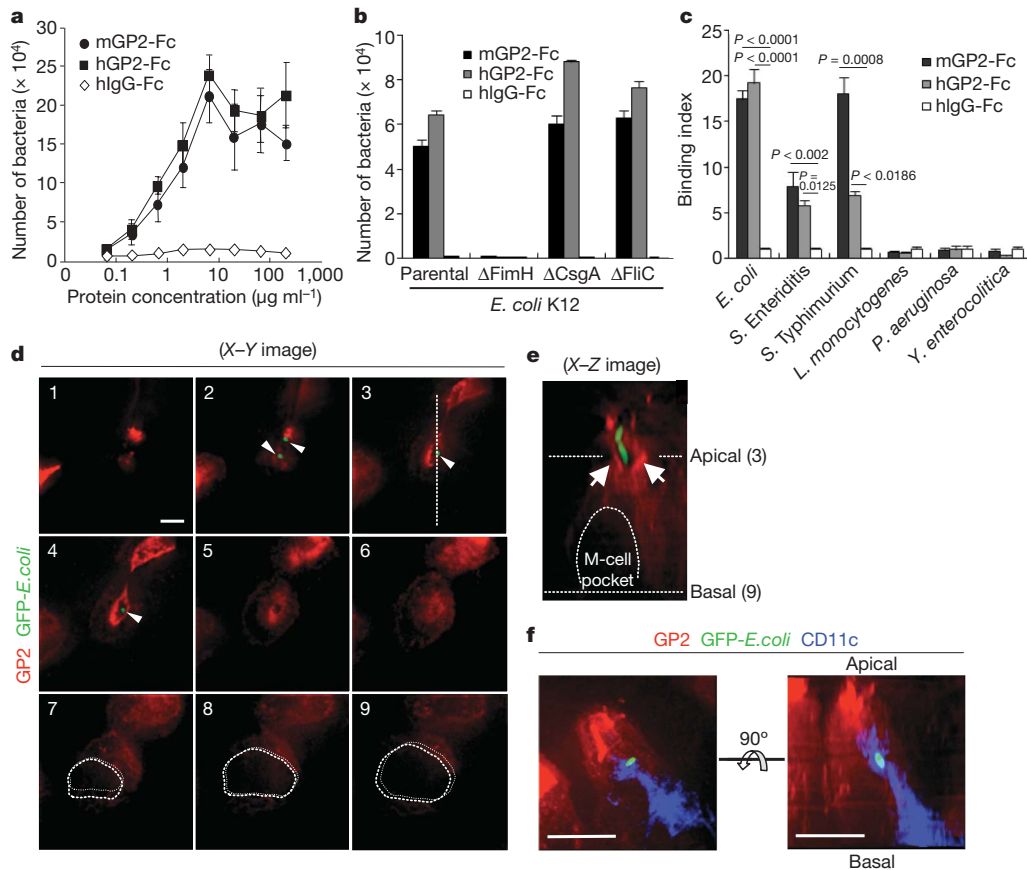
**Figure 2 | GP2 binds to FimH-expressing Gram-negative bacteria. a,** The *in vitro* binding assay was performed with *E. coli* and recombinant GP2-Fc and control hIgG-Fc proteins. Error bars indicate s.e.m. (*n* = 3). **b,** The GP2-binding capacity of *E. coli* was further examined using the parental K12 strain and *FimH*, *CsgA* or *FliC* knockout strains (ΔFimH, ΔCsgA or ΔFliC, respectively). Error bars indicate s.e.m. (*n* = 3). **c,** Interaction of recombinant GP2-Fc proteins with various commensal and pathogenic enterobacteria was examined by the *in vitro* binding assay. The binding index represents the fold change in the number of bacteria in hGP2-Fc or mGP2-Fc relative to that in control Fc. Error bars indicate s.e.m. (one-way analysis of variance followed by Tukey's test; *n* = 3 for all tested groups). **d,** After the ligated intestinal loop assay with GFP-*E. coli* (green), the whole-mount specimen of Peyer's patch was stained for GP2 (red), and analysed by

deconvolution microscopy. Serial *X–Y* sections of whole-mount staining of an M cell from apical (image 1) to basal (image 9) domains are shown. Arrowheads indicate GFP-*E. coli* just being internalized into the M cell. The M-cell pocket is indicated by a dotted circular line (images 7–9). **e,** The *X–Z* image of an M cell at the position indicated by a dotted line in image 3 is shown. Arrows indicate GP2 densely clustered around GFP-*E. coli*. For three-dimensional imaging, see Supplementary Movie 2. **f,** After the ligated intestinal loop assay, the Peyer's patch was stained for GP2 (red) and a DC marker CD11c (blue). In this image, the GFP-*E. coli* was transcytosed by an M cell from the apical surface to the M-cell pocket, into which a DC has migrated to capture the bacterium. Left, top view. Right, side view. For three-dimensional imaging, see Supplementary Movie 4.

*Y. enterocolitica*, which gains entry into M cells through binding to β1-integrin[22], between *Gp2*[+/+] and *Gp2*[−/−] mice (Fig. 3a, right). Taken together, these data underscore the importance of the binding of GP2 and FimH in the transcytosis of type-I-piliated bacteria by M cells.

It has been thought that the antigens transcytosed by M cells are transferred to immature DCs accumulating in the subepithelial dome (SED) region beneath the FAE[6]. The antigen-primed DCs would then undergo maturation and migrate to the T-cell-rich region in Peyer's patches for antigen presentation and subsequent initiation of antigen-specific mucosal immunity including IgA[+] B-cell formation[23]. We explored whether bacterial antigens transcytosed by means of the GP2-dependent pathway are captured by DCs. After the intestinal loop assay, we observed DCs capturing GFP-*E. coli* in the M-cell pocket, the microdomain formed by invagination of the basolateral plasma membrane (Fig. 2f and Supplementary Movie 4). Quantitative analysis showed that more than 90% of the translocated bacteria were captured by DCs (Supplementary Fig. 10).

We also examined whether GP2-dependent bacterial transcytosis contributes to the initiation of the antigen-specific immune response in Peyer's patches. We employed an established oral immunization model using recombinant *S.* Typhimurium expressing fragment C of

tetanus toxoid (r*Salmonella*-ToxC) as an antigen[24]. We here developed the *FimH*-deletion mutant strain of r*Salmonella*-ToxC. These bacteria grew normally but lost the ability to bind GP2 recombinant protein (Fig. 4a), consistent with FimH-deficient *E. coli* (Fig. 2b). Correspondingly, a deficiency of FimH greatly lowered bacterial entry into Peyer's patches and the mesenteric lymph node (MLN) (Fig. 4b). As a consequence, the FimH-deficient strain induced a weak antigen-specific helper-T-cell (T_H) response in Peyer's patches (Fig. 4c) in comparison with its parental strain, resulting in attenuated antigen-specific faecal IgA and serum IgG production (Fig. 4d, e). Similarly, *Gp2*[−/−] mice failed to induce antigen-specific T_H-cell and antibody immune responses after oral immunization with FimH[+] r*Salmonella*-ToxC (Fig. 3b–d and Supplementary Fig. 11). This was not due to a defect in the immunological function of *Gp2*[−/−] mice, because splenic tetanus toxoid (TT)-specific T_H response was normally induced in these mice after systemic immunization (Supplementary Fig. 12). These observations demonstrate that GP2-dependent bacterial transcytosis is involved in immunosurveillance for FimH[+] bacteria in the intestinal lumen.

The identification of GP2 expression by M cells from Peyer's patches and its functional characterization support the hypothesis that M cells express bacterial pattern-recognition molecules on their
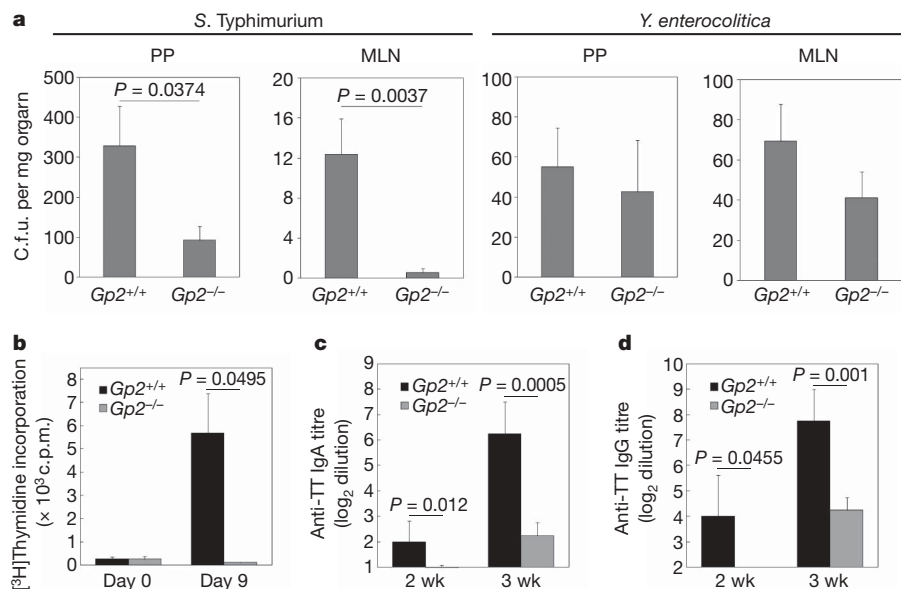
**Figure 3 | GP2 is important for the induction of an antigen-specific mucosal immune response.** **a,** $Gp2^{+/+}$ or $Gp2^{-/-}$ C57BL/6 mice were orally infected with $5 \times 10^7$ c.f.u. of *S.* Typhimurium or $10^9$ c.f.u. of *Y. enterocolitica.* Peyer's patches and MLN were collected 24 h after the infection to prepare tissue homogenates. The colonies of culturable bacteria in the tissue homogenates were counted and normalized to the weight of tissue samples, and are shown as colony-forming units. Data from two independent experiments were pooled and are shown as means and s.e.m. (Mann–Whitney *U*-test, $n = 6$). **b–d,** $Gp2^{+/+}$ or $Gp2^{-/-}$ C57BL/6 mice were orally immunized with $5 \times 10^7$ c.f.u. of r*Salmonella*-ToxC (FimH$^+$ strain). **b,** CD4 T cells were prepared from Peyer's patches 9 days after oral immunization, and were co-cultured for 4 days with irradiated splenic antigen-presenting cells obtained from C57BL/6 mice in the presence of TT to examine the proliferative response. Data are shown as means and s.d. ($n = 3$, Mann–Whitney *U*-test) and are representative of three independent experiments. **c, d,** The presence of TT-specific IgA in faeces (**c**) and IgG in serum (**d**) at 2 and 3 weeks after oral immunization was evaluated by ELISA ($n = 4$; Student's *t*-test). Data are shown as means and s.d. and are representative of four independent experiments.

apical surface that efficiently bind luminal microorganisms. Moreover, GP2-mediated transcytosis is necessary for the initiation of antigen-specific mucosal immune responses against type-I-piliated bacteria (Supplementary Fig. 1). Recent studies demonstrate that the translocation of *Salmonella* into Peyer's patches is required for the initiation of antigen-specific IgA responses[25,26], whereas DC-mediated delivery of the bacteria from intestinal lamina propria to the MLN and the spleen is responsible for inducing an antigen-specific IgG response. Our data showed that GP2 or FimH deficiency moderately, but significantly, decreased the antigen-specific IgG response after oral immunization with r*Salmonella*-ToxC (Figs 3d and 4e), most probably as a result of attenuated bacterial translocation to the MLN (Figs 3a and 4b). These observations raise the possibility that the entry of

type-I-piliated bacteria into M cells by means of the GP2–FimH system could also contribute to bacterial delivery to the MLN. An alternative, but not mutually exclusive, interpretation is that GP2 might also have a function in bacterial translocation to intestinal lamina propria. Given that the soluble form of GP2 is abundantly released from the exocrine pancreas into the gastrointestinal tract[13,27], the soluble GP2 might 'opsonize' luminal type-1-piliated bacteria and mediate efficient bacterial translocation at the lamina propria. Further study will be necessary to verify this speculation.

Our data therefore support the biological significance of antigen transcytosis by M cells in the front-line defence against enteric microbes by inducing mucosal immune responses. Future clarification of the GP2-mediated transcytotic mechanisms will define the
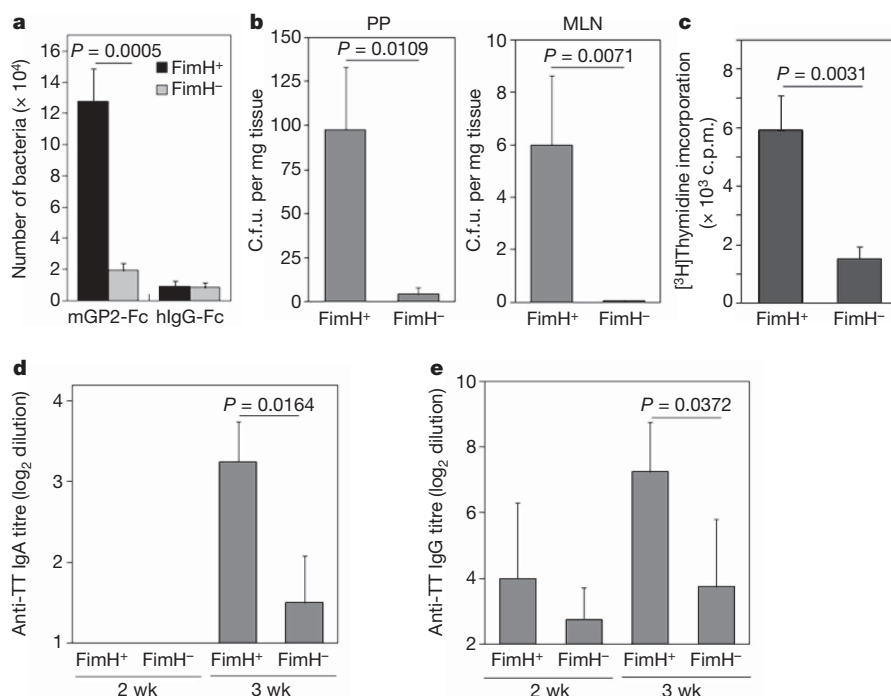


**Figure 4 | Bacterial translocation into Peyer's patches and subsequent antigen-specific mucosal immune response markedly decreased in FimH-deficinet *Salmonella*.** **a,** The *FimH* gene knockout (FimH$^-$) strain of r*Salmonella*-ToxC was developed as described in Methods. The *in vitro* binding assay was performed with parental (FimH$^+$) or FimH$^-$ r*Salmonella*-ToxC in microtitre plate immobilized with $5 \mu g ml^{-1}$ mGP2-Fc or control hIgG-Fc protein. Data are shown as means and s.d. ($n = 3$; Student's *t*-test). **b–e,** C57BL/6 mice were fed with $5 \times 10^7$ c.f.u. of FimH$^+$ or FimH$^-$ r*Salmonella*-ToxC. After 24 h, bacterial translocation to Peyer's patches and MLN was examined (**b**, $n = 5$; Mann–Whitney *U*-test). CD4 T cells were prepared from Peyer's patches 9 days after oral immunization, and were co-cultured for 4 days with irradiated splenic antigen-presenting cells obtained from the mice in the presence of TT to examine the proliferative response (**c**, $n = 3$; Student's *t*-test). The presence of TT-specific IgA in faeces (**d**) and IgG in serum (**e**) at 2 and 3 weeks after oral immunization were evaluated by ELISA ($n = 4$; Mann–Whitney *U*-test). Data are shown as means and s.d. and are representative of three independent experiments.

molecular basis of antigen transport by M cells and provide a new approach for 'M-cell-targeted' vaccination protocols.

## METHODS SUMMARY

**Animals.** BALB/c and C57BL/6 mice were purchased from CLEA Japan. GP2-deficient mice were generated as described previously[28,29], and were backcrossed onto a C57BL/6 genetic background. Mice were maintained under specific pathogen-free conditions in the animal facilities of RIKEN and the University of Tokyo until use in experiments. All animal experiments were approved by the Animal Research Committees of RIKEN Yokohama Research Institute and the Institute of Medical Science, the University of Tokyo.

***In vitro* bacterial binding assay.** Recombinant mGP2-Fc, hGP2-Fc or control Fc proteins were immobilized in 96-well polystyrene microtitre plates by incubation overnight at 4 °C. After washing, the microtitre wells were incubated with 1% BSA in PBS for 2 h for blocking, and then incubated for 2 h with $10^6$ colony-forming units (c.f.u.) of various bacteria in 1% BSA in PBS at 20–25 °C. After washing five times with sterile PBS, genomic DNA was extracted from bound bacteria with a QIAamp DNA kit (Qiagen). The copy number of 16S ribosomal RNA (rRNA) genes was determined by real-time qPCR with SYBR Premix Ex Taq (TaKaRa) and specific primers for universal 16S rRNA gene[30].

**Ligated intestinal loop assay.** Mice were anaesthetized with avertin and kept warm on a 37 °C warming pad during the assay. GFP-*E. coli* or GFP-*E. coli* (*FimH*−) ($10^8$ c.f.u.) or 50 μg ml$^{-1}$ anti-mouse GP2 mAb (clone 2F11-C3) were injected into the ligated intestinal loop. After incubation for 30–60 min, the mice were killed and Peyer's patches were excised from the intestine. Specimens were observed with a DM-IRE2 confocal laser scanning microscope (Leica Microsystems) or a DeltaVision Restoration deconvolution microscope (Applied Precision).

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 August; accepted 23 September 2009.

1. Cerutti, A. & Rescigno, M. The biology of intestinal immunoglobulin A responses. *Immunity* **28**, 740–750 (2008).
2. Fagarasan, S. *et al.* Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora. *Science* **298**, 1424–1427 (2002).
3. Peterson, D. A., McNulty, N. P., Guruge, J. L. & Gordon, J. I. IgA response to symbiotic bacteria as a mediator of gut homeostasis. *Cell Host Microbe* **2**, 328–339 (2007).
4. Craig, S. W. & Cebra, J. J. Peyer's patches: an enriched source of precursors for IgA-producing immunocytes in the rabbit. *J. Exp. Med.* **134**, 188–200 (1971).
5. Owen, R. L. & Jones, A. L. Epithelial cell specialization within human Peyer's patches: an ultrastructural study of intestinal lymphoid follicles. *Gastroenterology* **66**, 189–203 (1974).
6. Neutra, M. R., Mantis, N. J. & Kraehenbuhl, J. P. Collaboration of epithelial cells with organized mucosal lymphoid tissues. *Nature Immunol.* **2**, 1004–1009 (2001).
7. Nochi, T. *et al.* A novel M cell-specific carbohydrate-targeted mucosal vaccine effectively induces antigen-specific immune responses. *J. Exp. Med.* **204**, 2789–2796 (2007).
8. Sirard, J. C., Niedergang, F. & Kraehenbuhl, J. P. Live attenuated *Salmonella*: a paradigm of mucosal vaccines. *Immunol. Rev.* **171**, 5–26 (1999).
9. Sansonetti, P. J. & Phalipon, A. M cells as ports of entry for enteroinvasive pathogens: mechanisms of interaction, consequences for the disease process. *Semin. Immunol.* **11**, 193–203 (1999).
10. Kraehenbuhl, J. P. & Neutra, M. R. Epithelial M cells: differentiation and function. *Annu. Rev. Cell Dev. Biol.* **16**, 301–332 (2000).
11. Hase, K. *et al.* Distinct gene expression profiles characterize cellular phenotypes of follicle-associated epithelium and M cells. *DNA Res.* **12**, 127–137 (2005).
12. Terahara, K. *et al.* Comprehensive gene expression profiling of Peyer's patch M cells, villous M-like cells, and intestinal epithelial cells. *J. Immunol.* **180**, 7840–7846 (2008).
13. Hoops, T. C. & Rindler, M. J. Isolation of the cDNA encoding glycoprotein-2 (GP-2), the major zymogen granule membrane protein. Homology to uromodulin/Tamm–Horsfall protein. *J. Biol. Chem.* **266**, 4257–4263 (1991).
14. Verbrugghe, P., Kujala, P., Waelput, W., Peters, P. J. & Cuvelier, C. A. Clusterin in human gut-associated lymphoid tissue, tonsils, and adenoids: localization to M cells and follicular dendritic cells. *Histochem. Cell Biol.* **129**, 311–320 (2008).
15. Fukuoka, S., Freedman, S. D., Yu, H., Sukhatme, V. P. & Scheele, G. A. GP-2/THP gene family encodes self-binding glycosylphosphatidylinositol-anchored proteins in apical secretory compartments of pancreas and kidney. *Proc. Natl Acad. Sci. USA* **89**, 1189–1193 (1992).
16. Mo, L. *et al.* Ablation of the Tamm–Horsfall protein gene increases susceptibility of mice to bladder colonization by type 1-fimbriated *Escherichia coli. Am. J. Physiol. Renal Physiol.* **286**, F795–F802 (2004).
17. Pizarro-Cerda, J. & Cossart, P. Bacterial adhesion and entry into host cells. *Cell* **124**, 715–727 (2006).
18. Yu, S. & Lowe, A. W. The pancreatic zymogen granule membrane protein, GP2, binds *Escherichia coli* type 1 fimbriae. *BMC Gastroenterol.* **9**, 58 (2009).
19. Ewen, S. W. *et al. Salmonella enterica* var Typhimurium and *Salmonella enterica* var Enteritidis express type 1 fimbriae in the rat *in vivo. FEMS Immunol. Med. Microbiol.* **18**, 185–192 (1997).
20. McDonald, D. *et al.* Recruitment of HIV and its receptors to dendritic cell–T cell junctions. *Science* **300**, 1295–1297 (2003).
21. Tanaka, Y. *et al.* T helper type 2 differentiation and intracellular trafficking of the interleukin 4 receptor-α subunit controlled by the Rac activator Dock2. *Nature Immunol.* **8**, 1067–1075 (2007).
22. Clark, M. A., Hirst, B. H. & Jepson, M. A. M-cell surface β1 integrin expression and invasin-mediated targeting of *Yersinia pseudotuberculosis* to mouse Peyer's patch M cells. *Infect. Immun.* **66**, 1237–1243 (1998).
23. Iwasaki, A. & Kelsall, B. L. Localization of distinct Peyer's patch dendritic cell subsets and their recruitment by chemokines macrophage inflammatory protein (MIP)-3α, MIP-3β, and secondary lymphoid organ chemokine. *J. Exp. Med.* **191**, 1381–1394 (2000).
24. VanCott, J. L. *et al.* Regulation of mucosal and systemic antibody responses by T helper cell subsets, macrophages, and derived cytokines following oral immunization with live recombinant *Salmonella. J. Immunol.* **156**, 1504–1514 (1996).
25. Martinoli, C., Chiavelli, A. & Rescigno, M. Entry route of *Salmonella typhimurium* directs the type of induced immune response. *Immunity* **27**, 975–984 (2007).
26. Hashizume, T. *et al.* Peyer's patches are required for intestinal immunoglobulin A responses to *Salmonella* spp. *Infect. Immun.* **76**, 927–934 (2008).
27. Rindler, M. J. & Hoops, T. C. The pancreatic membrane protein GP-2 localizes specifically to secretory granules and is shed into the pancreatic juice as a protein aggregate. *Eur. J. Cell Biol.* **53**, 154–163 (1990).
28. Kobayashi, K., Yanagihara, K., Ishiguro, K. & Fukuoka, S. GP2/THP gene family of self-binding, GPI-anchored proteins forms a cluster at chromosome 7F1 region in mouse genome. *Biochem. Biophys. Res. Commun.* **322**, 659–664 (2004).
29. Yu, S., Michie, S. A. & Lowe, A. W. Absence of the major zymogen granule membrane protein, GP2, does not affect pancreatic morphology or secretion. *J. Biol. Chem.* **279**, 50274–50279 (2004).
30. Greisen, K., Loeffelholz, M., Purohit, A. & Leong, D. PCR primers and probes for the 16S rRNA gene of most species of pathogenic bacteria, including bacteria found in cerebrospinal fluid. *J. Clin. Microbiol.* **32**, 335–351 (1994).

## METHODS

**Quantitative PCR.** Total RNA was prepared from FAE and villus epithelium as described previously[11]. Mouse *Gp2* gene expression was examined by real-time qPCR with SYBR Premix Ex Taq (TaKaRa) and specific primers, mGP2 5′-GATA CTGCACAGACCCCTCCA-3′ (forward) and 5′-GCAGTTCCGGTCATTGAGGT A-3′ (reverse), as described previously[31]. The mouse *Gp2* expression level was normalized to that of the glyceraldehyde-3-phosphate dehydrogenase gene in each tissue, and data are expressed as fold change in FAE compared with that in villus epithelium.

**cDNA clones and plasmid construction.** To obtain constructs for fusion proteins of mouse (m) or human (h) GP2 with the Fc segment of human IgG1, cDNAs corresponding to the amino-terminal region of GP2 lacking the transmembrane domain were amplified by PCR with a RIKEN FANTOM clone 2310037I18 or human FAE, respectively, as template. Primers used were as follows: mGP2, 5′-GAAGATCTACCATGAAAAGGATGGTGGGTT-3′ (forward) and 5′-CCGCT CGAGTGTAGTGTGGGGAGTCCCC-3′ (reverse); hGP2, 5′-CGCGGATCCACC ATGGAAAGGATGGTGGGC-3′ (forward) and 5′-CCGCTCGAGTCCATTCAT GACACCGGGAGA-3′ (reverse). cDNA fragments were inserted into the *Bam*HI/*Xho*I cloning sites of a pcDNA3 expression vector (Invitrogen) containing a fragment encoding the Fc segment of human IgG1 to obtain mGP2-Fc and hGP2-Fc. For *in situ* hybridization, an N-terminal fragment of mouse and human GP2 (GP2-N) was prepared by using mouse or human GP2-Fc/pcDNA3 vectors and the following primers: mGP2-N, 5′-GAAGATCTCGATTCATCAGTTGCATGGTG-3′ (forward) and 5′-CCGCTCGAGGAAACCCATACCTCCCAGCA-3′ (reverse); hGP2-N, 5′-CGCGGATCCGCGATAAAAACATGAGCGGC-3′ (forward) and 5′-CCGCTCGAGCTCTCCAGAATGTTCCTGCAG-3′ (reverse). The cDNA fragments obtained were inserted into the *Bam*HI/*Xho*I cloning sites of the pcDNA3 expression vector.

**In situ hybridization.** Digoxigenin-labelled RNA probes were prepared by *in vitro* transcription with SP6 RNA polymerase (Roche), using mouse and human GP2-N/ pcDNA3 plasmid vectors digested with *Bam*HI as templates. Paraffin-embedded tissue sections prepared from murine Peyer's patches were deparaffinized, treated with proteinase K and hybridized for 16 h with 100 ng ml$^{-1}$ digoxigenin-labelled RNA probes at 60 °C. Specific binding was detected by incubation with alkaline phosphatase-conjugated anti-digoxigenin antibody (Roche Diagnostics) for 2 h at room temperature, and revealed with purple alkaline phosphatase substrate overnight (Genostaff). The sections were counterstained with UEA-1 or nuclear fast red. For *in situ* hybridization of hGP2, endoscopic biopsies or surgical sections were obtained from non-inflamed terminal ileum containing lymphoid follicles from patients who had previously given informed consent. Biopsy samples were fixed with 4% buffered formalin and hybridized with the hGP2-specific antisense probe or control sense probe as described above for mouse samples. These studies were approved by the Committee on Human Subjects in RIKEN and Tokyo Women's Medical University.

**Recombinant protein preparation.** Human embryonic kidney (HEK293T) cells were transiently transfected with the Fc-fusion expression vectors prepared as above and cultured for 7–10 days. Recombinant proteins secreted into the culture supernatant were purified with a HiTrap protein A HP affinity column (GE Healthcare). The recombinant fusion proteins mGP2(ΔTM)-Fc and hGP2(ΔTM)-Fc were used in the assay for bacterial binding and for mAb preparation.

**Development of GP2-specific antibodies.** Rabbit polyclonal antibodies specific for mGP2 were generated in response to a keyhole-limpet haemocyanin-conjugated 16-amino-acid peptide (CSEELGEYHVYKLQGT) corresponding to residues 154–169 of mouse GP2. Rat anti-mouse GP2 IgG2a (clones 2F11-C3 and 1F5-H9) and mouse anti-human GP2 IgG1 (clone 3G7-H9) mAbs were raised by immunization with recombinant mouse or human GP2-Fc fusion proteins, respectively.

**Immunoblotting.** HeLa cells transfected with full-length mGP2/pcDNA3 plasmid or tissues from 12-week-old BALB/c mice were homogenized in ice-cold lysis buffer containing 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.1% Triton X-100 and a protease inhibitor cocktail (Roche). Each homogenate was centrifuged at 10,000$g$ for 10 min at 4 °C. The supernatants were separated by SDS–PAGE and then transferred to an Immobilon-P membrane (Millipore). GP2 protein was probed with anti-mGP2 polyclonal antiserum or pre-immune serum. Specific binding of the primary antibody was detected with horseradish-peroxidase-conjugated secondary antibody followed by Immobilon Western HRP Substrate (Millipore).

**Immunofluorescent staining.** To detect mouse GP2, paraffin sections of Peyer's patches, isolated lymphoid follicles or caecal or colonic patches were dehydrated and treated with anti-mouse GP2 mAbs described above. The binding of primary Abs was revealed with the TSA fluorescence system (PerkinElmer) in accordance with the manufacturer's instructions. The sections were further stained with rhodamine-labelled UEA-1 (Vector Laboratories) and counterstained with 4,6-diamidino-2-phenylindole. Because specimens of human Peyer's patches showed autofluorescence, specific binding of the primary antibody was revealed with the DAB substrate chromogen system (Dako).

For whole-mount staining, mouse Peyer's patches were excised from the small intestine and washed vigorously in ice-cold PBS, then fixed and permeabilized with a Cytofix/Cytoperm kit (BD Bioscience). The whole-mount specimens of Peyer's patches were then stained with anti-GP2 mAb (2F11-C3) followed by secondary antibodies conjugated with Alexa488 or Alexa594. In some experiments the specimens were further stained with rhodamine-labelled UEA-1 and Alexa633-labelled phalloidin. The specimens were embedded in a 30% solution of glycerol in PBS, and analysed with a DM-IRE2 confocal laser scanning microscope (Leica Microsystems) and a DeltaVision Restoration deconvolution microscope (Applied Precision).

**Immunoelectron microscopy.** Excised tissues were fixed by immersion in 4% paraformaldehyde/0.1% glutaraldehyde. Immunoelectron microscopy was performed on cryo-ultrathin sections as described previously[32].

**Preparation of GFP-expressing *E. coli* and *S.* Typhimurium.** *E. coli* K-12, its *FimH*-deficient mutant, and *S.* Typhimurium were transformed with the bacterial GFP expression vector pKKGFP[33] to generate GFP-*E. coli* and GFP-*E. coli* ΔFimH, respectively. The GFP–*S.* Typhimurium[33] was provided by J. P. Kraehenbuhl.

**Preparation of FimH-deletion mutant of r*Salmonella*-ToxC.** The *fimH* deletion mutants of *Salmonella* Typhimurium harbouring toxC were constructed by using the method and plasmids of refs 34, 35 with slight modification. In brief, *S.* Typhimurium harbouring pACBSR was transformed with PCR product, which was created on pKD4 with primers S-fimH-H1 (5′-CCTGTATCCGTCCGGCG TCATAAAAGGAAAAATAGAGATGGTGTAGGCTGGAGCTGCTTC-3′) and S-fimH-H2 (5′-TAGCGATGAAAACGCGCCGAAGGATCATTATGCCTCCTT ACCATATGAATATCCTCCTTA-3′). After growth in Luria–Bertani medium containing arabinose, kanamycin-resistant transformants were selected.

**Immunization.** The procedure of oral immunization with r*Salmonella*-ToxC is described elsewhere[24]. In brief, $Gp2^{+/+}$ or $Gp2^{-/-}$ mice were fed with $5 \times 10^7$ c.f.u. of the parental or FimH-deletion mutant strain of r*Salmonella*-ToxC ($aroA^-$, $aroD^-$) in 200 μl of PBS. Alternatively, mice received an intraperitoneal injection with $5 \times 10^5$ c.f.u. of r*Salmonella*-ToxC for systemic immunization. To evaluate antigen-specific T-cell response, CD4 T cells purified from Peyer's patches and splenic cell suspension by magnetic cell separation (MACS; Miltenyi Biotec) with microbead-conjugated anti-CD4 were co-cultured with irradiated splenocytes from C57BL/6 mice in the presence of 100 μg ml$^{-1}$ TT (Biken Foundation) for 3 days. [$^3$H]Thymidine (Amersham) were added to culture media 16 h before harvest to measure T-cell proliferation. The antibody titre specific for TT or *Salmonella* membrane fraction in serum and faecal extracts was measured by ELISA as described previously[36,37].

**CFU assay after oral infection.** *S.* Typhimurium carrying a nalidixic acid resistance gene (χ3306)[38] was a gift from H. Matsui. $Gp2^{+/+}$ or $Gp2^{-/-}$ mice were fed with $5 \times 10^7$ c.f.u. of *S.* Typhimurium χ3306 or $10^9$ c.f.u. of *Y. enterocolitica* WA (ATCC 27729)[39]. After 24 h, Peyer's patches and ileal tissue were dissected and incubated at room temperature in sterile PBS containing 100 μg ml$^{-1}$ gentamicin for 30 min with gentle shaking. The tissues were weighed and homogenized in sterile PBS. The MLN was also removed and homogenized. The homogenates were serially diluted in sterile PBS and plated on a Luria–Bertani agar plate containing 25 μg ml$^{-1}$ nalidixic acid or a *Yersinia*-selective agar plate (Oxoid) to determine colony-forming units.

**Statistical analyses.** Differences between two or more groups were analysed by Student's *t*-test or a one-way analysis of variance followed by Tukey's test, respectively. When variances were not homogeneous, the data were analysed by Mann–Whitney *U*-test or Kruskal–Wallis ANOVA followed by a Scheffé test. All statements of significant differences show a 5% level of probability.

31. Hase, K. *et al.* The membrane-bound chemokine CXCL16 expressed on follicle-associated epithelium and M cells mediates lympho-epithelial interaction in GALT. *J. Immunol.* **176**, 43–51 (2006).
32. Waguri, S. & Komatsu, M. Biochemical and morphological detection of inclusion bodies in autophagy-deficient mice. *Methods Enzymol.* **453**, 181–196 (2009).
33. Hopkins, S. A., Niedergang, F., Corthesy-Theulaz, I. E. & Kraehenbuhl, J. P. A recombinant *Salmonella typhimurium* vaccine strain is taken up within murine Peyer's patch dendritic cells. *Cell. Microbiol.* **2**, 59–68 (2000).
34. Herring, C. D., Glasner, J. D. & Blattner, F. R. Gene replacement without selection: regulated suppression of amber mutations in *Escherichia coli*. *Gene* **311**, 153–163 (2003).
35. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
36. Jackson, R. J. *et al.* Optimizing oral vaccines: induction of systemic and mucosal B-cell and antibody responses to tetanus toxoid by use of cholera toxin as an adjuvant. *Infect. Immun.* **61**, 4272–4279 (1993).
37. Yamamoto, M. *et al.* Alternate mucosal immune system: organized Peyer's patches are not required for IgA responses in the gastrointestinal tract. *J. Immunol.* **164**, 5184–5191 (2000).
38. Gulig, P. A., Doyle, T. J., Hughes, J. A. & Matsui, H. Analysis of host cells associated with the Spv-mediated increased intracellular growth rate of *Salmonella typhimurium* in mice. *Infect. Immun.* **66**, 2471–2485 (1998).
39. Carter, P. B. & Collins, F. M. Experimental *Yersinia enterocolitica* infection in mice: kinetics of growth. *Infect. Immun.* **9**, 851–857 (1974).

# Cohesin acetylation speeds the replication fork

Marie-Emilie Terret[1], Rebecca Sherwood[1], Sadia Rahman[1], Jun Qin[2] & Prasad V. Jallepalli[1]

Cohesin not only links sister chromatids but also inhibits the transcriptional machinery's interaction with and movement along chromatin[1–6]. In contrast, replication forks must traverse such cohesin-associated obstructions to duplicate the entire genome in S phase. How this occurs is unknown. Through single-molecule analysis, we demonstrate that the replication factor C (RFC)–CTF18 clamp loader (RFC$^{CTF18}$)[1,7] controls the velocity, spacing and restart activity of replication forks in human cells and is required for robust acetylation of cohesin's SMC3 subunit and sister chromatid cohesion. Unexpectedly, we discovered that cohesin acetylation itself is a central determinant of fork processivity, as slow-moving replication forks were found in cells lacking the Eco1-related acetyltransferases ESCO1 or ESCO2 (refs 8–10) (including those derived from Roberts' syndrome patients, in whom *ESCO2* is biallelically mutated[11]) and in cells expressing a form of SMC3 that cannot be acetylated. This defect was a consequence of cohesin's hyperstable interaction with two regulatory cofactors, WAPL and PDS5A (refs 12, 13); removal of either cofactor allowed forks to progress rapidly without ESCO1, ESCO2, or RFC$^{CTF18}$. Our results show a novel mechanism for clamp-loader-dependent fork progression, mediated by the post-translational modification and structural remodelling of the cohesin ring. Loss of this regulatory mechanism leads to the spontaneous accrual of DNA damage and may contribute to the abnormalities of the Roberts' syndrome cohesinopathy.

The ring-shaped cohesin complex not only catenates sister DNAs, but also acts as a structural barrier to other chromatin-based transactions. For example, in budding and fission yeast cohesin slides to the 3′ ends of convergently arrayed genes and prevents readthrough transcription by RNA polymerase II (refs 2, 14). In mammals, cohesin accumulates at insulators and boundary elements and blocks activation of adjacent promoters[4,5]. These sites occur about once per 25 kilobases (kb) in the human and mouse genomes, an interval several times smaller than the average somatic replicon[15], and are occupied by cohesin throughout the cell cycle[4,5]. Cohesin can also bind regions between insulators in a distributive manner and pair sister chromatids through these sites[5]. Hence, nearly all forks must replicate through at least one cohesin-associated region during S phase. In addition to mediating genome duplication, fork passage is thought to lead to entrapment of nascent DNA strands within the cohesin ring, thereby generating functional cohesion between sister chromatids[1]. Cohesion establishment also requires acetylation of cohesin's SMC3 subunit by Eco1-related acetyltransferases[8–10], but how this modification affects cohesin's interaction with the replication fork remains unclear.

In addition to Eco1, cohesion establishment is aided by the replication factor C–CTF18–CTF8–DCC1 (hereafter RFC$^{CTF18}$) complex, which loads the polymerase processivity clamp proliferating cell nuclear antigen (PCNA) onto replication forks *in vitro* and *in vivo*[1,7]. Nevertheless, physical assays of replication fork dynamics have not uncovered this enzyme's contribution during a normal S

phase, nor is it known whether RFC$^{CTF18}$ is important for SMC3 acetylation. To address these questions, human retinal pigment epithelial cells were subjected to two rounds of adeno-associated virus (AAV)-mediated homologous recombination, such that exon 2 of the *DCC1* (also called *DSCC1*) locus was either flanked by *loxP* sites or deleted outright (Supplementary Fig. 1). To inactivate RFC$^{CTF18}$, *DCC1$^{flox/\Delta}$* cells were infected with an adenovirus expressing Cre recombinase (AdCre). The resulting cells lacked detectable DCC1 protein and had markedly reduced levels of CTF18, but normal amounts of all other RFC subunits (Fig. 1a). Whereas RFC$^{CTF18}$-deficient yeast grow robustly, *DCC1$^{\Delta/\Delta}$* clones could not be isolated
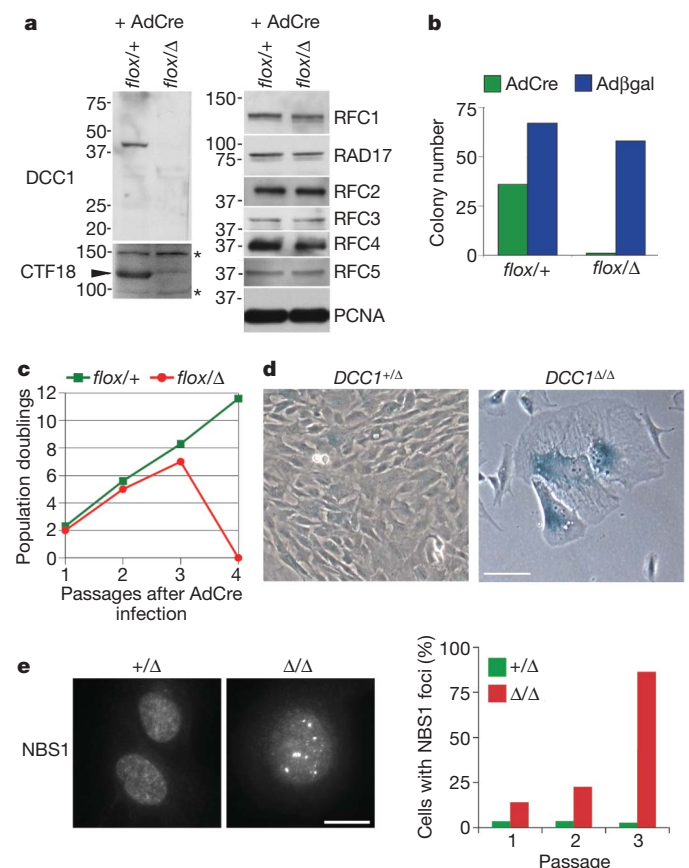


**Figure 1 | The RFC$^{CTF18}$ complex is required to prevent accumulation of endogenous DNA damage and terminal senescence. a**, Homozygous deletion of *DCC1* selectively inactivates RFC$^{CTF18}$. **b**, Cells were infected with AdCre or Adβgal and scored for colony formation. All *DCC1$^{flox/\Delta}$*-derived clones retained the parental genotype. **c**, Cells were serially passaged to monitor proliferation. Senescent *DCC1$^{\Delta/\Delta}$* cells at passage 3 failed to attach and survive after replating for passage 4, yielding a terminal population doubling value of zero. **d**, *DCC1*-null cells undergo senescence. Scale bar, 50 μm. **e**, *DCC1$^{\Delta/\Delta}$* cells accumulate NBS1-positive DNA damage foci. Scale bar, 10 μm.

[1]Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10065, USA. [2]Center for Molecular Discovery, Verna and Marrs McLean, Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas 77030, USA.

via limiting dilution (Fig. 1b), indicating that this complex is essential in mammals. To understand this requirement, $DCC1^{flox/\Delta}$ cells were infected with AdCre en masse and monitored during serial passage. Within seven divisions, $DCC1^{\Delta/\Delta}$ cells ceased to proliferate and exhibited hallmarks of premature senescence (increased size, flattened morphology and expression of senescence-associated β-galactosidase; Fig. 1c, d). Senescence is commonly triggered by genotoxic stresses that activate the DNA damage response[16]; consistently, $DCC1^{\Delta/\Delta}$ cells accumulated high levels of NBS1 foci during cell cycle exit (Fig. 1e). Unlike RFC[RAD17]-deficient cells[17], RFC[CTF18]-deficient cells activated CHK1 and imposed the G2/M checkpoint in response to genotoxic stress (Supplementary Fig. 2).

Unstable replication forks are a prime source of DNA damage in both prokaryotes and eukaryotes[18]. As an initial test of the contribution of RFC[CTF18] to fork stability, replication forks in newly generated $DCC1^{+/\Delta}$ and $DCC1^{\Delta/\Delta}$ cells were transiently stalled with hydroxyurea (which depletes nucleotide pools) or aphidicolin (which inhibits replicative polymerases). RFC[CTF18]-deficient cells were hypersensitive to both replication inhibitors, as evidenced by their accelerated transit into senescence (Fig. 2a). To obtain a more precise view of replication dynamics, single-molecule analyses were performed on extended DNA fibres. Cells were sequentially pulse-labelled with the halogenated nucleosides iododeoxyuridine (IdU) and chlorodeoxyuridine (CldU), after which genomic DNA was gently extracted and stretched on glass slides, stained with IdU- and CldU-specific antibodies, and visualized microscopically (Fig. 2b). Unlike other assays for DNA synthesis, this technique affords specific and quantitative information on origin firing, replication fork movement and fork recovery throughout the genome, on the basis of the examination of large numbers of individually labelled forks. Three effects of RFC[CTF18] deficiency were evident in these assays. First, the velocity of individual forks decreased by more than 50% ($0.7 \pm 0.2$ against $1.7 \pm 0.4$ kb min$^{-1}$ in $DCC1^{+/\Delta}$ cells; mean ± s.d., Fig. 2c), indicating a highly penetrant defect in replisome progression. Second, the interval between forks decreased significantly (from $38 \pm 7.6$ to $18 \pm 4.5$ kb; Fig. 2d), consistent with firing of dormant origins that usually are replicated passively[19]. Third, replication forks appeared to arrest or collapse at an elevated rate, as reflected by an excess of tracks labelled exclusively with IdU (Fig. 2e). As a further test, we measured the efficiency of fork recovery from an aphidicolin block, imposed between the IdU and CldU pulses. Over 30% of forks failed to restart under these conditions (Fig. 2e), thus explaining the sensitivity of $DCC1^{\Delta/\Delta}$ cells to replication inhibitors.

The role of the RFC[CTF18] complex in cohesion establishment was also examined. As senescing cells are excluded from metaphase spreads, we abrogated this p53- and Rb-mediated response via expression of human papillomavirus proteins E6 and E7 (ref. 20). Bypassing senescence did not alter the requirement for RFC[CTF18] in sustaining high fork velocity, origin dormancy and replication restart (Supplementary Fig. 3), but did reveal precocious chromatid separation in over 40% of $DCC1^{\Delta/\Delta}$ cells, as well as structural aberrations (Fig. 3a, b). A similar mixture of cohesion loss and chromosome fragility occurs in cells expressing non-acetylatable SMC3 (ref. 8), suggesting that this modification may depend on RFC[CTF18]. To test this idea,
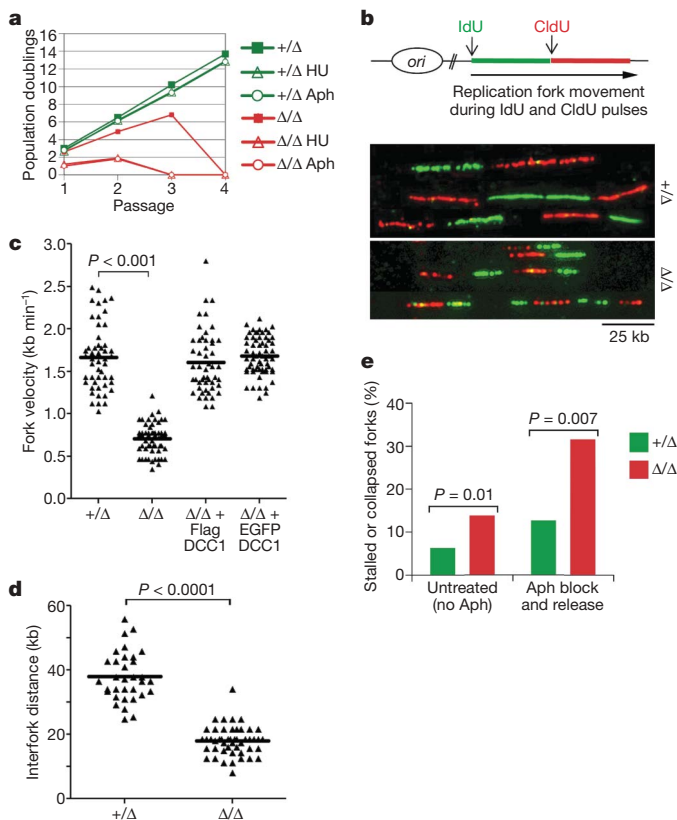
**Figure 2 | RFC[CTF18] controls global replication dynamics and promotes fork reactivation after genotoxic stress. a**, Cells were transiently exposed to hydroxyurea (HU) or aphidicolin (Aph) and monitored for growth during subsequent passages. **b**, Single-molecule analysis of replication fork dynamics in $DCC1^{+/\Delta}$ and $DCC1^{\Delta/\Delta}$ cells. **c, d**, Fork velocities (**c**) and distances between forks (**d**) were measured from 50 tracks per sample. **e**, Cells were either labelled with IdU and CldU as above (left columns), or treated with 10 µM aphidicolin for 2 h after the IdU pulse, then washed and pulsed with CldU (right columns).
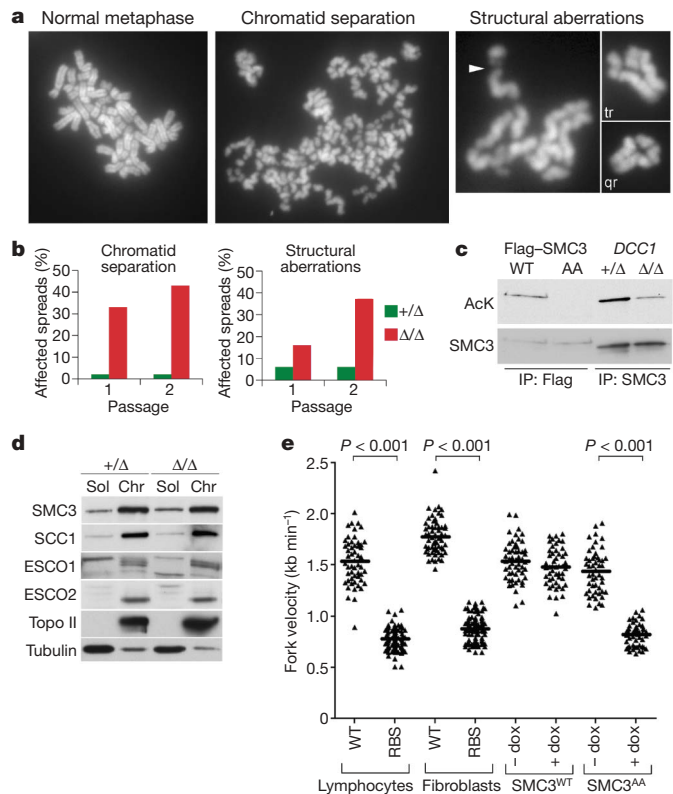
**Figure 3 | SMC3 acetylation and replication fork progression are interdependent. a, b**, Karyotypic analysis of E6- and E7-expressing $DCC1^{+/\Delta}$ and $DCC1^{\Delta/\Delta}$ cells. Chromatid gaps (arrowhead), triradial (tr), and quadriradial (qr) chromosomes are highlighted. **c**, Cells in **a** were synchronized in S phase with thymidine. Chromatin-associated SMC3 was immunopurified and blotted for acetylated lysine or total SMC3. Wild-type (WT) and non-acetylatable (AA) SMC3 variants[8] were included as controls. **d**, Soluble (Sol) and chromatin (Chr) fractions from S-phase cells were analysed by immunoblotting (Topo II, topoisomerase II). **e**, Fork velocities were determined in cells from normal donors and RBS patients, and in cells expressing SMC3[WT] or SMC3[AA] from a doxycycline-inducible promoter (dox).

cohesin was immunopurified and analysed by quantitative immuno-blotting with antibodies specific for acetylated lysine. SMC3 acetylation was reduced by 70% in $DCC1^{\Delta/\Delta}$ cells (Fig. 3c), despite proper targeting of cohesin and the Eco1-related acetyltransferases ESCO1 and ESCO2 to chromatin (Fig. 3d). Thus, RFC$^{CTF18}$ promotes cohesion at a later step than the Cdc7 kinase and other pre-replicative complex components, which are required for cohesin's deposition onto interphase chromosomes in $Xenopus$[21]. These data are compatible with a model wherein ESCO1 and ESCO2 are stimulated to acetylate cohesin through physical interactions with RFC$^{CTF18}$-loaded PCNA[1,22].

Our results show that, in the absence of RFC$^{CTF18}$, replication fork processivity and SMC3 acetylation are both significantly diminished. We next asked whether these processes occur independently down-stream of RFC$^{CTF18}$ or are themselves linked. To this end, we analysed replication dynamics in cells derived from patients with Roberts' syndrome/SC phocomelia (RBS), a severe hereditary disorder in which $ESCO2$ is biallelically mutated, resulting in weakened cohesion between sister centromeres[11]. Fork velocities were markedly reduced in cells from two unrelated RBS patients (Fig. 3e) and accompanied by spontaneous accumulation of DNA damage (Supplementary Fig. 4). Fork slowing was also observed in HeLa cells depleted of ESCO2 or ESCO1 via RNA interference (RNAi; see Fig. 4e). To test the contribution of cohesin acetylation more specifically, we analysed replication forks in cells expressing wild-type (SMC3$^{WT}$) or non-acetylatable SMC3 (SMC3$^{AA}$) from a tetracycline-regulated promoter. Expression of SMC3$^{AA}$ alone reduced fork speeds from $1.5 \pm 0.2$ to $0.8 \pm 0.1$ kb min$^{-1}$ (Fig. 3e), indicating that cohesin acetylation per se is required for processive DNA synthesis. In contrast, the expression of cohesin-regulated genes was insensitive to these manipulations (Supplementary Fig. 5 and Supplementary Table 1). We conclude that the acetylated form of cohesin specifically regulates replication fork progression, but is not required for transcriptional insulation.

How SMC3 acetylation affects cohesin biochemically is not well understood, but genetic evidence indicates that this modification acts in opposition to WAPL and PDS5, which together bind cohesin and regulate its association with chromosomes[10,12,13,23–26]. Strikingly, non-acetylatable human SMC3 co-purified with increased amounts of both WAPL and PDS5A (Fig. 4a and Supplementary Fig. 6). Furthermore, this interaction was negatively correlated with SMC3 acetylation in synchronized cells (Fig. 4b and Supplementary Fig. 7). To determine whether the excess stability of this interaction inhibits fork movement, RNAi was used to deplete WAPL and PDS5A in cells expressing SMC3$^{AA}$ (Fig. 4c, d) or concomitantly depleted of ESCO1 or ESCO2 (Fig. 4e and Supplementary Fig. 8). Notably, downregulating WAPL or PDS5A rescued fork progression in all instances, but had no effect on its own, showing that ring opening by the WAPL–PDS5A complex is probably not generally required for replication forks to move through cohesin-binding sites. PDS5A depletion also allowed forks in $DCC1$-null cells to move at normal rates (Fig. 4f–g). We conclude that the RFC$^{CTF18}$ complex speeds chromosomal replication mainly by stimulating the acetylation-dependent remodelling of cohesin rings, rather than by activating DNA polymerases directly, at least in the presence of the replicative RFC.

Our analysis provides new and unexpected insights into the relationship between DNA replication and cohesion establishment. Previous studies did not detect an S phase delay in yeast $ctf18$ or $eco1$ mutants, leading to the conclusion that chromosomes are replicated in a similar fashion whether or not cohesion is established between nascent sister DNAs. In contrast, by applying sensitive single-molecule assays, we find that replication dynamics in human cells are highly attuned to the acetylation of SMC3 and possibly other substrates by ESCO1 and ESCO2. Notably, the fork's response to deficiencies in cohesin acetylation is indistinguishable from the consequences of physically blocking the fork by other routes (for example, nucleotide depletion, polymerase inhibition or DNA alkylation), suggesting that a primary role of this modification is to
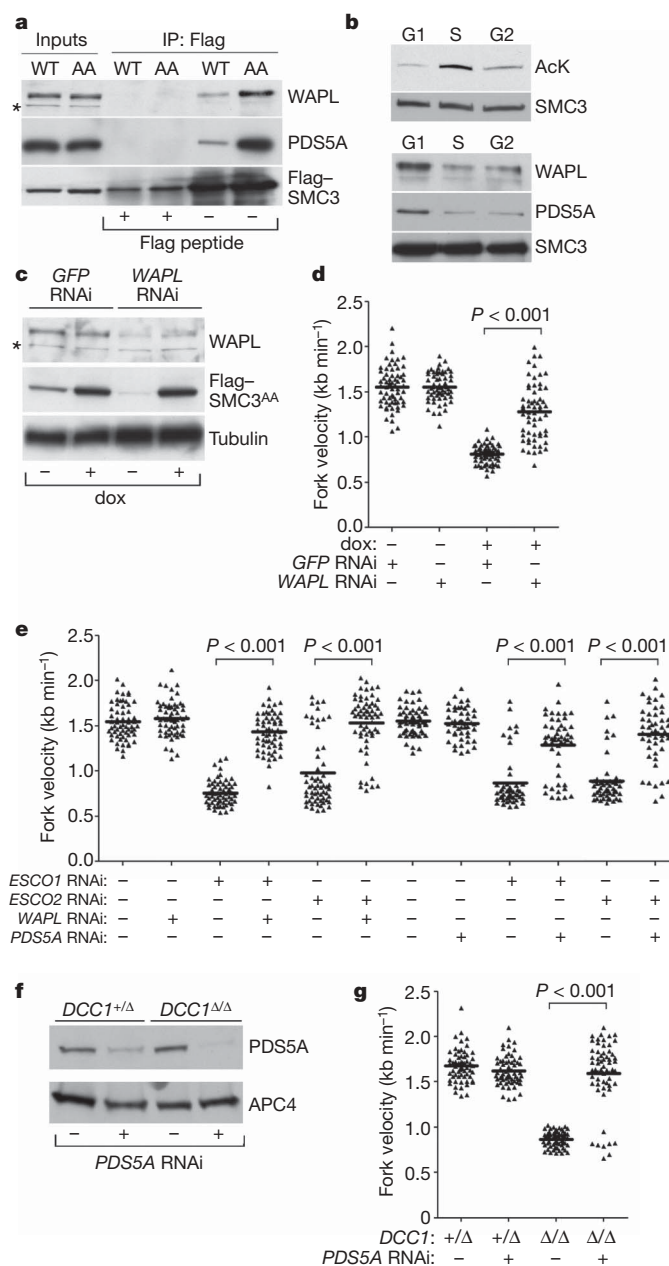
**Figure 4 | Processive fork movement requires acetylation-mediated dissociation of PDS5A and WAPL. a**, SMC3 variants were immunopurified and examined for WAPL and PDS5A binding. **b**, Cells were synchronized by serum withdrawal (G1 phase) or thymidine block and release (S and G2). Acetylation of chromatin-bound SMC3 (top) and association with WAPL and PDS5A (bottom) were determined as in Figs 3c and 4a. **c, d**, WAPL depletion rescues fork progression in SMC3$^{AA}$-expressing cells.
**e**, Downregulating WAPL or PDS5A reverses fork slowing in ESCO1- and ESCO2-depleted cells. **f, g**, PDS5A depletion rescues fork progression in the absence of RFC$^{CTF18}$.

switch cohesin from a configuration that obstructs the fork to one that permits its advancement. Furthermore, our genetic and biochemical data imply that this switch involves cohesin's regulatory cofactors PDS5A and WAPL, rather than an intrinsic biophysical property of acetylated cohesin itself. To account for these observations, we propose that WAPL and PDS5A prevent the replication fork from sliding through the ring and/or interfere with cohesin's stable embrace of the nascent sister DNAs (Supplementary Fig. 9). The possibility of an occlusive conformation is already suggested by PDS5's bivalent interactions with opposite sides of the ring[27], as well as flexion of the SMC1 and SMC3 coiled-coils[28]. To neutralize this, the fork-associated RFC$^{CTF18}$ complexes stimulate ESCO1 and ESCO2 to

acetylate SMC3, causing the WAPL–PDS5A complex to dissociate. Cohesin could then adopt an open conformation that allows the fork to advance while stably trapping the nascent sister DNAs within the ring. This model would also explain the 'anti-establishment' activities ascribed to WAPL and PDS5A[10,23,25,26], because hindrance of fork passage through the cohesin ring would not only slow chain elongation, but also interfere with entrapment of sister chromatids. In addition to generating cohesion, entrapment may contribute positively to fork processivity by facilitating sister-chromatid-dependent pathways of replication restart[29].

Weakened centromere cohesion is a diagnostic hallmark of RBS[11]. However, we show here that this disorder is also associated with widespread slowing of the replication fork, presumably due to unopposed action of the WAPL–PDS5A complex on cohesin. Whether this chronic alteration in replication dynamics contributes to the abnormalities in RBS patients is unclear. However, DNA replication in general is known to be epigenetically coordinated with development. Whereas active origins are closely spaced during the initial cell cycles after fertilization, their density falls more than tenfold during embryogenesis, resulting in a dramatic expansion of replicon size[15]. By impeding fork progression and reactivating dormant origins, defects in cohesion establishment may disrupt this epigenetic switch, causing replication stress and premature senescence or apoptosis of progenitor cells during development. Further work with *in vivo* models of RBS will be needed to test this hypothesis. Reducing fork velocity also results in the packaging of chromosomal DNAs into smaller than normal chromatin loops[19]. This effect has been suggested to involve changes in the replication-dependent modification of factors associated with matrix attachment regions and boundary elements, a proposal compatible with the known properties of cohesin[4,5]. Its acetylation may therefore be required not only to produce cohesion locally, but also to shape mammalian genomes more globally.

## METHODS SUMMARY

To mutate the *DCC1* locus, hTERT-RPE1 cells were infected with AAV vectors with homology to sequences flanking exon 2. Clones with correct gene replacements were identified by PCR and Southern blotting. Purified adenoviruses expressing Cre recombinase or β-galactosidase were obtained from Baylor College of Medicine Vector Development Laboratory and applied at a multiplicity of infection of 200. Senescence bypass was accomplished through retroviral expression of human papillomavirus E6 and E7 oncoproteins. RNAi was initiated by transfection of synthetic siRNA duplexes via Oligofectamine (Invitrogen), followed by analysis of protein depletion, replication fork movement, and sister chromatid cohesion 48 to 72 h later. Techniques for preparing and analysing IdU- and CldU-labelled DNA fibres were adapted from procedures published previously[30]. Images were acquired on a TE2000 microscope fitted with a ×100 1.4NA oil objective and analysed using the NIS-Elements software package (Nikon Instruments).

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Lengronne, A. *et al.* Establishment of sister chromatid cohesion at the *S.cerevisiae* replication fork. *Mol. Cell* **23**, 787–799 (2006).
2. Gullerova, M. & Proudfoot, N. J. Cohesin complex promotes transcriptional termination between convergent genes in *S.pombe. Cell* **132**, 983–995 (2008).
3. Rubio, E. D. *et al.* CTCF physically links cohesin to chromatin. *Proc. Natl Acad. Sci. USA* **105**, 8309–8314 (2008).
4. Parelho, V. *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
5. Wendt, K. S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
6. Stedman, W. *et al.* Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.* **27**, 654–666 (2008).
7. Bermudez, V. P. *et al.* The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. *Proc. Natl Acad. Sci. USA* **100**, 10237–10242 (2003).
8. Zhang, J. *et al.* Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. *Mol. Cell* **31**, 143–151 (2008).
9. Unal, E. *et al.* A molecular determinant for the establishment of sister chromatid cohesion. *Science* **321**, 566–569 (2008).
10. Ben-Shahar, T. R. *et al.* Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. *Science* **321**, 563–566 (2008).
11. Vega, H. *et al.* Roberts syndrome is caused by mutations in *ESCO2*, a homolog of yeast *ECO1* that is essential for the establishment of sister chromatid cohesion. *Nature Genet.* **37**, 468–470 (2005).
12. Gandhi, R., Gillespie, P. J. & Hirano, T. Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr. Biol.* **16**, 2406–2417 (2006).
13. Kueng, S. *et al.* Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127**, 955–967 (2006).
14. Lengronne, A. *et al.* Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573–578 (2004).
15. Walter, J. & Newport, J. W. Regulation of replicon size in *Xenopus* egg extracts. *Science* **275**, 993–995 (1997).
16. Deng, Y., Chan, S. S. & Chang, S. Telomere dysfunction and tumour suppression: the senescence connection. *Nature Rev. Cancer* **8**, 450–458 (2008).
17. Wang, X. *et al.* Rad17 phosphorylation is required for claspin recruitment and Chk1 activation in response to replication stress. *Mol. Cell* **23**, 331–341 (2006).
18. Heller, R. C. & Marians, K. J. Replisome assembly and the direct restart of stalled replication forks. *Nature Rev. Mol. Cell Biol.* **7**, 932–943 (2006).
19. Courbet, S. *et al.* Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature* **455**, 557–560 (2008).
20. Shay, J. W., Pereira-Smith, O. M. & Wright, W. E. A role for both RB and p53 in the regulation of human cellular senescence. *Exp. Cell Res.* **196**, 33–39 (1991).
21. Takahashi, T. S., Basu, A., Bermudez, V., Hurwitz, J. & Walter, J. C. Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in *Xenopus* egg extracts. *Genes Dev.* **22**, 1894–1905 (2008).
22. Moldovan, G. L., Pfander, B. & Jentsch, S. PCNA controls establishment of sister chromatid cohesion during S phase. *Mol. Cell* **23**, 723–732 (2006).
23. Tanaka, K., Hao, Z., Kai, M. & Okayama, H. Establishment and maintenance of sister chromatid cohesion in fission yeast by a unique mechanism. *EMBO J.* **20**, 5779–5790 (2001).
24. Losada, A., Yokochi, T. & Hirano, T. Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and *Xenopus* egg extracts. *J. Cell Sci.* **118**, 2133–2141 (2005).
25. Rowland, B. D. *et al.* Building sister chromatid cohesion: Smc3 acetylation counteracts an antiestablishment activity. *Mol. Cell* **33**, 763–774 (2009).
26. Sutani, T., Kawaguchi, T., Kanno, R., Itoh, T. & Shirahige, K. Budding yeast Wpl1(Rad61)-Pds5 complex counteracts sister chromatid cohesion-establishing reaction. *Curr. Biol.* **19**, 492–497 (2009).
27. Mc Intyre, J. *et al. In vivo* analysis of cohesin architecture using FRET in the budding yeast *Saccharomyces cerevisiae. EMBO J.* **26**, 3783–3793 (2007).
28. Sakai, A., Hizume, K., Sutani, T., Takeyasu, K. & Yanagida, M. Condensin but not cohesin SMC heterodimer induces DNA reannealing through protein–protein assembly. *EMBO J.* **22**, 2764–2775 (2003).
29. Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nature Rev. Mol. Cell Biol.* **9**, 297–308 (2008).
30. Jackson, D. A. & Pombo, A. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J. Cell Biol.* **140**, 1285–1295 (1998).

**Author Contributions** M.-E.T. and P.V.J. designed experiments, M.-E.T., R.S. and S.R. performed experiments, J.Q. contributed reagents, M.-E.T., R.S., S.R. and P.V.J. analysed the data, and M.-E.T. and P.V.J. wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.V.J. (jallepap@mskcc.org).

## METHODS

**Gene targeting and RNA interference.** Sequences surrounding exon 2 of human *DCC1* were amplified from BAC clone RP11-188E9 with *PfuTurbo* (Stratagene) and cloned on either side of a central *FRT-neo^R-FRT-loxP* cassette. A second *loxP* site was introduced upstream of exon 2, and the entire insert was then moved into pAAV as a NotI fragment. The final construct was fully sequenced to verify its integrity. Preparation of infectious AAV particles, gene-targeting procedures and Southern blotting were performed as described previously[31,32]. Excision of the *FRT*-flanked *neo^R* cassette was accomplished by transfection with pCAGGS-FLPe (Open Biosystems), followed by limiting dilution to recover G418-sensitive colonies. During the second round of gene targeting, we identified polymorphisms within the 5′ and 3′ homology arms that biased recombination towards the already targeted allele. A vector isogenic to untargeted allele was constructed and used to delete exon 2. Synthetic 21-nucleotide siRNA duplexes (*GFP*, target GGCTA CGTCCAGGAGCGCA; *WAPL*, target CGGACTACCCTTAGCACAA; *ESCO1*, targets GGACAAAGCTACATGATAG (*ESCO1* #1) and GAGAATAAATTTCC AGGTT (*ESCO1* #2); *ESCO2*, target TAAGTCCACTGTCTATCCA; *SCC1*, target TGGAAGATCTCCTAACTAA; *PDS5A*, target TTCTTCCTCAGGAACCCCATT) were obtained from Dharmacon and transfected using Oligofectamine. Mock RNAi was performed by omitting siRNAs from the transfection mixture.

**Cell culture and drug treatments.** hTERT-RPE1 cells were grown in Dulbecco's modified essential medium (DMEM) F-12 supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin. HeLa and HEK293 Flp T-Rex cells were grown in DMEM with 10% FBS and 1% penicillin–streptomycin. Flag-SMC3 expression was induced with 1 μg ml$^{-1}$ doxycycline. Human lymphoblastoid cell lines (GM07524 and GM20466) and primary fetal fibroblasts (GM06112 and GM21872) were grown in RPMI 1640 with 15% FBS and 1% penicillin–streptomycin and DMEM with 15% FBS and 1% penicillin–streptomycin, respectively. Nocodazole (0.2 μg ml$^{-1}$), hydroxyurea (1 mM), adriamycin (0.5 μM), thymidine (2.5 mM), caffeine (5 mM) and aphidicolin (0.3–30 μM) were also used.

**Retroviral transduction.** Epitope-tagged forms of DCC1 were cloned into pQCXIN (Clontech). A pLXSN clone containing human papillomavirus serotype 16 E6 and E7 was obtained from D. Galloway. Vectors were transfected into Phoenix cells together with a VSV-G envelope expression plasmid. Virion-containing supernatants were supplemented with 10 μg ml$^{-1}$ polybrene and applied to target cells for 15 h. Transductants were selected with 0.4 mg ml$^{-1}$ G418.

**Pulse-labelling of replication forks and DNA fibre analysis.** Exponentially growing cells were labelled with 50 μM IdU for 20 min, then 50 μM CldU for 20 min. Extended DNA fibres were prepared using a modification of a procedure described previously[30]. Briefly, labelled cells were trypsinized and resuspended in ice-cold PBS at 1 × 10$^6$ cells ml$^{-1}$. Two microlitres of this suspension was spotted onto a clean glass slide and lysed with 10 μl of spreading buffer (0.5% SDS in 200 mM Tris-HCl (pH 7.4) and 50 mM EDTA). After 6 min, the slides were tilted at 15° to horizontal, allowing the DNA to spread. Slides were air-dried, fixed in methanol and acetic acid (3:1) for 2 min, and refrigerated overnight before immunolabelling. DNA was denatured with 2.5 M HCl for 30 min at room temperature. Slides were rinsed three times in PBS and blocked in PBS + 0.1% Triton X-100 (PBS-T) + 10% goat serum for 1 h at room temperature. Rat anti-BrdU (Abcam ab6326, 1:100) and mouse anti-BrdU (Becton Dickinson 347580, 1:100) were then applied to detect CldU and IdU, respectively. After a 1-h incubation, slides were washed three times in PBS and stained with Alexa Fluor 488-labelled goat anti-mouse IgG1 antibody and Alexa Fluor 594-labelled goat anti-rat antibody (Invitrogen, 1:350 each). Slides were mounted in Prolong Plus (Invitrogen) and held at 4 °C overnight. Replication tracks were imaged on a Nikon TE2000 microscope fitted with a ×100 1.4NA oil objective and measured using NIS Elements software. Restart assays were conducted exactly as above, except that cells were treated with 10 μM aphidicolin for 2 h after the IdU pulse. The drug was then washed out and the CldU pulse was administered. Continuity of DNA fibres was verified by staining with an anti-DNA antibody (Millipore MAB3034, 1:300) and an appropriate secondary antibody (Alexa 647-labelled goat anti-mouse IgG2a (Invitrogen), 1:350). Fork-velocity data sets were evaluated for statistical significance using one-way analysis of variance (ANOVA) and Tukey's post hoc test.

**Population doubling assays and immunofluorescence microscopy.** Cells (15–20% confluent) were infected with AdCre or Adβgal at a multiplicity of infection of 200. On day 3 after infection, cells were counted with a haemocytometer and replated at a 1:16 ratio. This point was defined as passage 0. On day 4, cells were treated with hydroxyurea or aphidicolin for 15 h (if indicated). On day 6, cells were again harvested and counted. This point was defined as passage 1. Two further rounds of subculturing were used to obtain cells at passages 2 and 3. Senescence-associated β-galactosidase staining was performed as described[33]. For immunofluorescence, cells were grown in chamber slides and fixed with 4% paraformaldehyde at room temperature or 100% methanol at −20 °C for

20 min. After permeabilization in PBS-T for 10 min, cells were blocked in PBS-T + 10% goat serum for 30 min, incubated with primary and Alexa 488- and 594-conjugated secondary antibodies for 1 to 2 h each, counterstained with 4,6-diamidino-2-phenylindole (DAPI), and mounted in Prolong Plus (Invitrogen).

**DSP cross-linking, immunoprecipitation and chromatin fractionation.** To stabilize labile protein complexes, dithiobis(succinimidyl propionate) (DSP) was added to the culture medium at 1 mg ml$^{-1}$ for 10 min at 37 °C. Extracts were prepared by lysing pellets on ice for 30 min in HB2 buffer (50 mM HEPES, pH 7.5, 0.5% NP-40, 10% glycerol, 100 mM NaCl, 10 mM Na pyrophosphate, 5 mM β-glycerophosphate, 50 mM NaF, 0.3 mM Na$_3$VO$_4$, 1 mM PMSF and 1× complete protease inhibitor cocktail), followed by sonication and centrifugation at 13,000*g* for 30 min at 4 °C. The extract (3 mg) was immunoprecipitated with Flag M2 antibody-agarose beads. Where indicated, a 50-fold excess of Flag peptide was added as a competitive inhibitor. Beads were washed three times with HB2 before elution and reversal of cross-links with Laemmli buffer.

To analyse acetylation of chromatin-bound SMC3, DSP was omitted, and cells were lysed in NETN buffer[8] supplemented with 10 mM sodium butyrate. After centrifugation the chromatin pellet was resuspended in the same buffer, sonicated, and recentrifuged. Soluble chromatin was immunoprecipitated with antibodies to SMC3 or Flag as indicated. For quantification, western blots were incubated with species-specific secondary antibodies conjugated to IRDye 680 or IRDye 800 and scanned on an Odyssey infrared imaging system (Li-Cor Biosciences). Chromatin fractionation in Fig. 3d was performed using the procedure of ref. 34.

**Antibodies.** The following antibodies were used: rabbit anti-DCC1 and anti-CTF18 (gift of J. Hurwitz); rabbit anti-ESCO1 and anti-ESCO2 (gift of H. Zou); rabbit anti-NBS1 (gift of J. Petrini); rat anti-BrdU (Abcam) and mouse anti-BrdU (Becton Dickinson); mouse anti-Flag (Sigma); mouse anti-RAD17, anti-p53 and anti-CHK1, rabbit anti-RFC1 and anti-PCNA (Santa Cruz); goat anti-RFC2, rabbit anti-RFC3, anti-RFC4, anti-RFC5, anti-PDS5A, anti-SCC1, anti-SMC3, anti-WAPL and anti-APC4 (Bethyl Laboratories); rabbit anti-acetyl-lysine (Abcam); mouse anti-topo II (Millipore); rabbit anti-pS345-CHK1 (Cell Signaling); rat anti-tubulin (Chemicon). A polyclonal antibody to ESCO1 was prepared by immunizing rabbits with a recombinant fragment (amino acids 1–454) and affinity-purified before use.

**Insulator assays and microarrays.** Insulator assays were performed as described previously[5]. Briefly, HeLa cells were transfected with SCC1, ESCO1, ESCO2 or control siRNAs and synchronized by a double thymidine block (14 h thymidine 2.5 mM, 8 h release, 16 h thymidine 2.5 mM, 6 h release). During the release from the first thymidine arrest the firefly luciferase reporter plasmid (pIHLIE or pIHLME) and a *Renilla* luciferase control plasmid were transfected in a ratio of 1:20 using Lipofectamine Plus (Invitrogen). After harvesting and lysis of the cells, the activities of both luciferases were detected using a Dual-Luciferase Reporter Assay System kit (Promega) and a Glomax 96 microplate luminometer (Promega). To control for transfection efficiency, the firefly luciferase activity was normalized against the *Renilla* luciferase activity. Each condition was analysed in triplicate. The reporter plasmids pIHLIE and pIHLME have been described[35].

For expression profiling of endogenous cohesin-regulated genes, total RNA was isolated from SMC3$^{AA}$ and SMC3$^{WT}$ cells. Two biological replicates were examined per condition. Quality of RNA was ensured before labelling by analysing 20–50 ng of each sample using the RNA 6000 NanoAssay and a Bioanalyser 2100 (Agilent). Samples with a 28S/18S ribosomal peak ratio of 1.8–2.0 were considered suitable for labelling. For samples meeting this standard, 200 ng of total RNA were labelled using the Illumina TotalPrep RNA Amplification kit (Ambion), according to manufacturer's instructions. Labelled and fragmented cRNA (3 μg) was hybridized to HumanRef8 arrays (Illumina). Microarray bioinformatics analyses were conducted with GeneSpring GX v.10.0 software (Agilent). A collection of 36 genes proximal to cohesin-binding sites and known to be transcriptionally regulated by cohesin[5] was interrogated. Relative expression of these genes in SMC3$^{AA}$ and SMC3$^{WT}$ cells was evaluated using unpaired *t*-test and Benjamini–Hochberg false discovery rate correction.

31. Papi, M., Berdougo, E., Randall, C. L., Ganguly, S. & Jallepalli, P. V. Multiple roles for separase auto-cleavage during the G2/M transition. *Nature Cell Biol.* **7**, 1029–1035 (2005).

32. Burkard, M. E. *et al.* Chemical genetics reveals the requirement for Polo-like kinase 1 activity in positioning RhoA and triggering cytokinesis in human cells. *Proc. Natl Acad. Sci. USA* **104**, 4383–4388 (2007).

33. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic *ras* provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593–602 (1997).

34. Méndez, J. & Stillman, B. Chromatin association of human origin recognition complex, Cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell. Biol.* **20**, 8602–8612 (2000).

35. Ishihara, K., Oshimura, M. & Nakao, M. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol. Cell* **23**, 733–742 (2006).

# Q&A

**Sheila Widnall** of the Massachusetts Institute of Technology (MIT), Cambridge, won the Arthur M. Bueche Award for expanding opportunities for women and minorities in engineering.

**What was your first big break?**

My uncle worked for a mining company in Colorado. He gave me a chunk of uranium and I used it, along with models of atoms, to explain radioactive decay. I won the science fair in Tacoma, Washington, as a high-school junior, which garnered attention from some MIT alumni and other mentors who encouraged me to go to school at MIT.

**Have you had a career-defining moment?**

I'm a fluid dynamicist. In this field, the greatest honour awarded is to have an instability, the forces behind complex flow dynamics, named after you. The Widnall instability is named for my work on how unstable waves develop in turbulent fluid flows.

**You have achieved many firsts. Which are you most proud of?**

I was the first woman faculty chair at MIT and the first woman to serve as president of the American Institute of Aeronautics and Astronautics, but being the first woman to serve as Secretary of the Air Force was a special honour. As someone who always wanted to fly every plane that the Air Force has, it was the type of incredible opportunity that makes my colleagues' eyes water.

**How would you describe your early opportunities?**

I have lived at the leading edge of a wave of increased sensitivity to the role of women in science and engineering. Science associations, MIT and even corporations were all working to increase the visibility of women. I benefited enormously from people stepping forward in my life to encourage me to take advantage of opportunities. It has been an incredible ride.

**What has been the biggest change for women during your working life?**

The numbers and the acceptance of women is definitely the biggest change. I was one of 20 women in a freshman class of 1,100. Now women make up almost half of the undergraduate population. The most significant change, though, is the sense of equality and teamwork among the students. It is a very different atmosphere from when I was in school.

**Is there a continued need to champion opportunities for women in engineering, or do they have parity?**

I think there is a continuing need. MIT has done very well. Women are the majority in half the engineering departments here, but that is not true everywhere in the country. And, in industry, women are still not in the highest levels of management. We still need to pay attention to that.

**What is your career philosophy?**

Aim high, the Air Force credo, has worked for me. ∎

Interview by Virginia Gewin

---

**POSTDOC JOURNAL**

# Nice to be appreciated

On 24 September, postdocs at Canadian universities celebrated the first annual National Postdoc Appreciation Day (NPAD). Designated by the US National Postdoctoral Association (NPA), the day recognizes and aims to raise awareness of postdocs' many research contributions.

Here in Canada, a handful of universities held NPAD events, including the University of Ottawa, where about 70 postdocs celebrated with a 'Happy NPAD' cake. At the University of Calgary, our postdoctoral association bought us a pizza lunch, over which about 50 postdocs spoke with our associate vice-president about postdoc issues on campus.

Cakes, pizzas and parties are a nice start. But with so few official postdoctoral associations at Canadian universities, and even fewer institutes recognizing NPAD, we have yet to raise our profile enough to advocate at provincial and federal levels. We do, fortunately, now have an organization in place. The recently established Canadian Association of Postdoctoral Scholars (CAPS) should be lobbying for nationwide improvements and standardized policies on work conditions, career development and professional support. CAPS is run by postdoc volunteers, and needs formal and federal support to sustain itself. Still, there is reason to believe we'll see improvement in the future.

The NPA is now organizing events for the 2010 NPAD. Perhaps at next year's party we'll have more to celebrate. ∎

**Julia Boughner is a postdoc in evolutionary developmental biology at the University of Calgary, Canada.**

---

# IN BRIEF

## UK engineering up

The number of students starting undergraduate engineering courses at UK colleges and universities in 2009 is way above last year's figures, according to a report issued by the Universities and Colleges Admissions Service (UCAS), which has been tracking these numbers since 1993. Aerospace engineering admissions rose by 20.7% to 2,016; mechanical by almost 12% to 5,757; electronic/electrical by 3.7% to 4,654; and civil by 3.2% to 4,432. Admissions to undergraduate physics programmes rose by 7% to 3,559. UCAS had earlier reported that applications for admission in all programmes were up for the year by 9.7% over 2008 numbers.

## Eagle eye on resources

Nine US universities are launching a publicly accessible network to connect scientists with information that is usually unavailable. The online network — dubbed 'eagle-i' by the project team — will give access to data held on the internal networks of institutions, which could include some of the country's top labs. Health researchers studying a particular group of people, for example, could access information on patient populations. The US$15-million pilot scheme is funded by the US National Center for Research Resources, which is also putting $12.2 million into funding the expansion of the cross-disciplinary collaboration network VIVO (see *Nature* 462, 123; 2009).

## Chinese science online

Aiming to provide an information and networking resource for scientists and others in China, the Chinese government has launched a new website, www.escience.gov.cn. The China S&T Resources Sharing Website — in Chinese only — includes science news, information about scientific instruments and equipment, and links to resources of use to Chinese scientists such as worldwide publications and data-sharing networks. Users can instant-message one another through a link on the site. The website, which is publicly accessible, is administered by China's ministry of science and technology, which says that the project is part of a move to build a national science-and-technology resource infrastructure.

# A BRIDGE TO SOMEWHERE

Graduate students who head abroad to study face any number of challenges if they hope to prosper. **Virginia Gewin** provides a study guide.

Delgerjargal Dugarjav doesn't mind travelling long distances to pursue an education. She left her native Mongolia to earn a master's degree in India and went on, in 2008, to start a doctorate in forest ecology at the University of Wisconsin-Madison. Dugarjav wanted to study how carbon stored in the boreal forests of her home country might help climate-change mitigation. That kind of graduate research was simply not feasible in Mongolia owing to a lack of resources and expertise.

She expected academic challenges, but she had not anticipated the intensity of the transition — the culture shock and, in particular, the heightened level of competition in the United States. "Foreign students have to try much harder than US students to do the same work," she says.

Students such as Dugarjav who choose to study abroad face formidable financial, social and academic challenges, including difficulty finding affordable lodging, adapting to a new language and facing competition from fellow students. The transition may be rocky but, with assistance, is eminently manageable.

Dugarjav is among a steady stream of science students pursuing international academic dreams. The two most popular destinations for these aspiring graduate scholars are the United States and the United Kingdom. The United States, for example,

welcomed 68,069 students from India, 53,047 from China and 24,697 from South Korea in 2007, according to the Institute of International Education (IIE), a non-profit organization based in New York City that promotes educational exchange.

In the United Kingdom, the number of foreign students almost doubled between 1998 and 2007 from 117,290 to 229,640, according to London-based Universities UK, which represents the country's higher-education institutions.

## Finding a good fit

A successful transition starts with finding the academic programmes that best suit one's specific needs. Students may be tempted to base their decision on an institution's general rankings, such as those conducted by *US News & World Report* or *Times Higher Education*. However, field rankings mean the most, especially for science and technology graduate programmes, says Peggy Blumenthal, executive vice-president and chief operating officer of the IIE. "Harvard and the University of California at Berkeley are good schools, but

it is important to ask if they are good in your chosen field of study," she says.

David Stevens, head of the International Office at University College London, suggests that identifying and establishing contact with potential supervisors at a school of choice can increase the success of an application. "When choosing a school, individuals should be ambitious, but also realistic," he says.

Contacting a professor or student at a university of interest can help students gauge their potential success there. Dugarjav first set her sights on the University of Wisconsin-Madison when she met a student from the university during a 2001 workshop in Mongolia on long-term ecological research. The University of Wisconsin-Madison forest ecology programme, she discovered, was a good fit with her interest in carbon cycling.

Several organizations make it their mission to provide insights into academia, culture and economics for students from abroad. The British Council — the United Kingdom's international organization for cultural relations and educational opportunities — has offices in 110 countries

**Delgerjargal Dugarjav: busy.**

around the globe. Staff can help students in those countries to explore different UK universities' specialities, seek scholarships and funding support, and learn about life in different UK cities. Likewise, EducationUSA, a network run by the US Department of State, has more than 450 offices operating in 170 countries. Its staff aims to inform prospective students about the diverse education opportunities at more than 400,000 US colleges and universities.

When choosing a university, students should also consider the cost, the climate, the size of the city, the local amenities and the cultural diversity, say university advisers. "We find that many foreign nationals value the cultural diversity found in all the major UK cities," says Catherine Marston, a policy adviser at Universities UK.

Metropolitan centres aren't the only ones to value cultural diversity. Purdue University in West Lafayette, Indiana, for example, has an 'international friendship programme' that matches community residents with new students. Some 175 students were matched last autumn. Purdue's Office of International Students and Scholars organizes a variety of social and cultural outings — from basketball games to Amish-country field trips — for the approximately 40% of Purdue's graduate students who come from abroad. "We try to give our international students a rich cultural experience," says programme director Michael Brzezinski.

### The funding factor

Finding a way to pay for a rich academic experience is the next hurdle. Many foreign graduate students secure funding through their host country or through a research or teaching fellowship from the department of the university they plan to attend. Purdue funds 70% of its students through this mechanism, which again highlights the benefit of establishing departmental contacts.

Partnership schemes designed to support individual nationalities, such as Commonwealth scholarships that aim to attract students from Singapore or Australia to the United Kingdom, or partnerships with companies looking to target specific research areas, also exist at most universities. But government-sponsored scholarships, such as the UK Chevening programmes or the US Fulbright programmes, are among the most prestigious awards open to foreign nationals. Dugarjav jumped at the chance to apply for a fully funded PhD through the International Fulbright Science and Technology Award. The programme gives out roughly 45 awards each year to outstanding foreign students in fields from neuroscience to astronomy. The three-year awards pay a monthly stipend plus tuition and fees, books, equipment, travel, research and professional conference allowances. After that, the universities

continue their support until the students complete their programmes. Currently, 40% of Fulbright scholars do science research.

But it's not just about academia. College advisers make it clear that admission committees are interested in a person's leadership qualities. "Scores and grades speak for themselves, but students need to make sure their personality, energy and passion for their research field are evident on the application," says Danielle Guichard-Ashbrook, associate dean for international students at the Massachusetts Institute of Technology in Cambridge. Foreign students, especially those in an area where progress is largely determined by test scores, may not appreciate the need to reveal their motivations, says Guichard-Ashbrook.

### Mixed feelings

Once at university, students should be prepared to experience a range of emotions. Exploring cultural differences makes the first few weeks in a new country exciting. But once the excitement ebbs, typically after about two months, many students become frustrated, both culturally and academically.

"All students go through culture shock — it's very normal," says Joanna Shearer, head of international student recruitment at Imperial



> "All students go through culture shock — it's very normal. It is important to keep a sense of humour."
> — Joanna Shearer

College London. She says it is important to keep a sense of humour, especially with respect to cultural differences. Her office, as with those at many universities, offers a tutorial of regional slang terms for incoming students as part of their orientation. "The students who are successful understand that there will be differences and will come with a curiosity to explore those differences," says John Greisberger, director of the International Center at the University of Michigan in Ann Arbor. He says good schools offer programmes that help students adjust culturally and in other ways.

John Pearson, director of Stanford University's Bechtel International Center in Palo Alto, California, has found that foreign students appreciate a clear explanation of the rights and responsibilities they must uphold as part of their foreign student status in the United States. For example, students must obtain a social security number and maintain an accurate home address in the university system in order to conform with federal regulations requiring universities to keep track of foreign students. (See 'Tips for success'.)

Nonetheless, students must find a way to balance their academic and social pursuits, notes Guichard-Ashbrook. "The most successful and happiest students are the ones who have a life outside the lab," she says. Martin Petricic, a Croatian PhD student in ocean engineering at the University of California at Berkeley, has continued his lifelong hobby of fencing in an effort to meet more people. Petricic views his time in the United States as an opportunity to make contacts. "These contacts may open doors to future research collaborations," he says.

For Dugarjav, making the most of her US graduate experience means finding a way to return to Mongolia to not only conduct research, but also to help educate the younger generation. "I would really like to use what I have learned," she says, "to improve the education system back home." ■

**Virginia Gewin is a science writer based in Portland, Oregon.**

---

## TIPS FOR SUCCESS

**DO** apply for a student visa as soon as possible. The visa process in both the United States and Britain, for example, has undergone revisions in recent years. Lengthy delays are no longer the norm but can occur.

**DON'T** assume that student visas are general and allow entry to any university at which you have been accepted. Student visas are site-specific. New US and UK regulations, designed to track foreign students more effectively, grant visas to one institution.

**DO** get involved with the student club dedicated to your particular nationality.

**DON'T** let that be your only connection to other students. Explore other ways to connect with fellow students.

**DO** seek academic help or tutoring if needed.

**DON'T** think that asking for help is a sign of weakness. International student offices provide a number of support programmes to help students be successful. **V.G.**

---

**Clarification**
The Feature 'Seeds of collaboration' (*Nature* **461**, 1158–1159; 2009) omitted that the Searle family topped up the US$5 million it gave to establish the Chicago Biomedical Consortium with a further $15 million. In addition, it should be made clear that the University of Chicago is the sole institution managing Argonne National Laboratory.

# An open letter

To any impressionable young school leavers who are considering joining the space corps.

**Martin Hayes**

If you are considering joining the space corps, I've got one word for you: don't.

Don't do it.

No, seriously now, don't even think about it.

Wait. What's that you say? *Why the hell not?* Yeah, I knew you'd say that. Okay wise guy, here goes nothing.

You know those adverts that you see during your favourite TV shows? They're all lies. That foxy waitress who is so impressed by the young hotshot who just graduated from the academy that her clothes just seem to fall off of their own volition — they think we're all morons. That handsome, well-built guy with the oiled biceps dodging the ship through the asteroid belt — never happened, would never happen, will never happen.

It's not glamorous, it's not exciting, it pays less than what I used to make as a welder and it will really screw up your social life. Still don't believe me? Huh, I figured as much.

Okay, so you probably think that life in the space corps is all thrills, spills and madcap adventures, gritting your teeth while you make split-second, life or death decisions. Yeah?

Wrong.

The truth is that it's supremely, monstrously, interminably boring. Even if you're picked for the induction course and you manage to trudge your way through all the mental training and assault courses, the personality tests and press-ups, it can still take years before you even set foot on a ship to begin your first mission. And believe me, that's when the fun really goes out of it.

Do you want to know the worst thing about waking up from a 19-month cryosleep?

The size of the bloody erection, that's what.

Seriously, you could hammer tent pegs into Europan sheet-ice with the thing. It's agony. On my last trip out it took five days for the bruising to fade away. And I had to sit down to take a leak. You try going to the ship's doctor and telling her that you've got acute thrombosis of the penis. There's a conversation you won't forget in a hurry.

And then there are all the other hassles of cryo-sleep. You haven't washed your face for a year and a half so you're riddled with zits, and if you've slept a bit funny on your arm you'll be walking like Quasimodo for the next three weeks — and the breath!

Plus, it's practically a given that at least one of your colleagues will be playing tonsil hockey with some kind of hideous face-sucking parasite. You'll try to take it off but it never works. It'll have acid for blood or something, as usual. We used to run a book on who would be the one sporting the alien balaclava but it's not worth

it any more, you can't get good odds. Last time out I was 3 to 1 on.

But the parasites aren't even the half of it.

If I told you about all the times my fellow crew members have been haunted by ethereal-yet-physical representations of their recently deceased wives, I'd be here all day. It's almost become a cliché at this point. For some guys, that's the only reason they go on missions anymore, to catch up with their dead wives, have a bit of a chat, let them know how the kids are getting on, all that.

And the nocturnal noises — just let me tell you this: it's no fun trying to get back to sleep when you've been woken at a quarter to four in the morning by the syrupy moans of your shipmate getting frisky with the ghost of his dead wife in the cabin next door. There isn't a bulkhead in the world thick enough to shield your ears from those high-pitched supernatural delights.

And everyone knows that the companies who run the ships are a bunch of mean sods. They never put proper lighting in the hallways, they're always dark and shadowy and a little bit on the scary side, and for some unknown reason there's always a leaking pipe somewhere, usually down in the cargo bay, constantly drip-drip-dripping water and driving everyone crazy.

Here's what I'd like to know, if they can build a spaceship the size of a small city and fly it through interstellar space at almost the speed of light, why can't they fix a leaking pipe? Eh? Well?

You should also forget all those exciting rumours you've heard about the aliens. Apart from the face-sucking-leech-beasts (who, despite their tendency to pounce on you when you least expect it, can at least be taught some basic domestic skills), we've found only one other species of intelligent life and they turned out to be the extra-solar equivalent of Morris Dancers — boring bastards to a man, and no, that's not sexist, they don't have females, they spontaneously reproduce in a little transparent bag that they carry around between their knees, and yes, that is as pretty as it sounds.

So, take my advice, stay on Earth. Get a job in a supermarket or something. It's less hassle, the money is better and at least I won't have to look at your sallow, boggle-eyed face over breakfast.

Yours sincerely,
Frank Ellison
Hydrogeological Survey Assistant
18th July, 2217
DS Research Ship *Ptolemy* — docked at Ganymede space port.

PS. I forgot to mention: never, on any account, and I mean never, trust the AI computer. It wants to kill you.

PPS. Seriously. Don't. ∎

**Martin Hayes lives in Arklow, a small town on the east coast of Ireland. He spends his days drinking strong tea and staring maniacally at a Crookes Radiometer. You can visit him at www.paroneiria.com.**
**Join the discussion of Futures in *Nature* at go.nature.com/QMAm2a**

JACEY

FUTURES